10/076905

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 8, 2005, 03:07:05 ; Search time 95 Seconds
(without alignments)
1126.574 Million cell updates/sec

Title: AAH26175
Perfect score: 1084
Sequence: 1 MKFKLHVNSARQYKDLWNMS.....PSPTSSTVITQAPSSNRPIV 209

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1084	100.0	209	2	Q8TAR1	Q8tar1 homo sapien
2	984	90.8	234	2	Q8BN75	Q8bn75 mus musculu
3	984	90.8	420	2	Q8CBR9	Q8cbr9 mus musculu
4	984	90.8	487	1	ATF2_HUMAN	P15336 homo sapien
5	984	90.8	487	1	ATF2_MOUSE	P16951 mus musculu
6	984	90.8	487	2	Q8CGB4	Q8cgb4 mus musculu
7	980	90.4	487	1	ATF2_RAT	Q00969 rattus norv
8	960	88.6	487	1	ATF2_CHICK	O93602 gallus gall
9	923	85.1	486	2	Q91576	Q91576 xenopus lae
10	719	66.3	440	2	Q68FE3	Q68fe3 mus musculu
11	571	52.7	413	2	Q8R0S1	Q8r0s1 mus musculu
12	555.5	51.2	494	1	ATF7_HUMAN	P17544 homo sapien
13	497	45.8	307	2	Q9UD83	Q9ud83 homo sapien
14	417	38.5	117	2	Q8IVR8	Q8ivr8 homo sapien
15	398.5	36.8	148	2	Q75N02	Q75n02 homo sapien
16	390.5	36.0	508	1	CRB5_HUMAN	Q02930 homo sapien
17	317.5	29.3	135	2	Q8BKN9	Q8bkn9 mus musculu
18	317.5	29.3	144	2	Q8BM42	Q8bm42 m mus muscu
19	208	19.2	144	2	Q96JT8	Q96jt8 homo sapien
20	167.5	15.5	850	2	Q9W0Z5	Q9w0z5 drosophila
21	133.5	12.3	593	2	Q8QGH4	Q8qgh4 brachydanio
22	122.5	11.3	503	2	Q9NAR8	Q9nar8 branchiosto
23	116.5	10.7	775	1	Z515_HUMAN	Q8nea6 homo sapien
24	115	10.6	183	2	Q9PVD6	Q9pvd6 brachydanio
25	115	10.6	442	2	O93311	O93311 brachydanio
26	115	10.6	565	1	SCR1_SCHPO	O14335 schizosacch
27	114.5	10.6	454	2	Q6CGR7	Q6cgr7 yarrowia li
28	114	10.5	780	2	Q9PWQ1	Q9pwq1 fugu rubrip
29	114	10.5	780	2	Q9YGM3	Q9ygm3 fugu rubrip
30	112	10.3	447	1	ZIC1_HUMAN	Q15915 homo sapien
31	110.5	10.2	1367	1	AMYH_YEAST	P08640 saccharomyc
32	110.5	10.2	1367	2	Q6LCS8	Q6lcs8 saccharomyc
33	110.5	10.2	1426	2	Q7QJK9	Q7qjk9 anopheles g
34	110	10.1	447	1	ZIC1_MOUSE	P46684 mus musculu
35	110	10.1	447	2	Q6PAK5	Q6pak5 mus musculu
36	110	10.1	447	2	Q9JKY2	Q9jky2 rattus norv
37	110	10.1	453	2	Q80Y18	Q80y18 mus musculu
38	110	10.1	881	2	Q69Z25	Q69z25 mus musculu
39	109.5	10.1	428	1	RSV1_SCHPO	Q9p7d9 schizosacch
40	109.5	10.1	674	2	Q80W63	Q80w63 mus musculu
41	109.5	10.1	675	1	MTF1_MOUSE	Q07243 mus musculu
42	109.5	10.1	675	2	Q9JJW8	Q9jjw8 mus musculu
43	109	10.1	568	2	Q8BMU0	Q8bmu0 mus musculu
44	109	10.1	791	2	Q9GRX2	Q9grx2 drosophila
45	109	10.1	791	2	Q9NFS1	Q9nfs1 drosophila

ALIGNMENTS

RESULT 1
Q8TAR1
ID Q8TAR1 PRELIMINARY; PRT; 209 AA.
AC Q8TAR1;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ATF2 protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC026175; AAH26175.1; -.
DR HSSP; P15336; 1BHI.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 1.
DR SMART; SM00355; ZnF_C2H2; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 1.
SQ SEQUENCE 209 AA; 23050 MW; A26AF07CA5D8D5B7 CRC64;

Query Match 100.0%; Score 1084; DB 2; Length 209;

```
  Best Local Similarity   100.0%;  Pred. No. 1.2e-71;
  Matches  209;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MKFKLHVNSARQYKDLWNMSDDKPFLCTAPGCGQRFTNEDHLAVHKHKHEMTLKFGPARN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKFKLHVNSARQYKDLWNMSDDKPFLCTAPGCGQRFTNEDHLAVHKHKHEMTLKFGPARN 60

Qy         61 DSVIVADQTPTPTRFLKNCEEVGLFNELASPFENEFKKASEDDIKKMPLDLSPLATPIIR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DSVIVADQTPTPTRFLKNCEEVGLFNELASPFENEFKKASEDDIKKMPLDLSPLATPIIR 120

Qy        121 SKIEEPSVVETTHQDSPLPHPESTTSDEKEVPLAQTAQPTSAIVRPASLQVPNVLLTSSD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SKIEEPSVVETTHQDSPLPHPESTTSDEKEVPLAQTAQPTSAIVRPASLQVPNVLLTSSD 180

Qy        181 SSVIIQQAVPSPTSSTVITQAPSSNRPIV 209
              |||||||||||||||||||||||||||||
Db        181 SSVIIQQAVPSPTSSTVITQAPSSNRPIV 209


RESULT 2
Q8BN75
ID   Q8BN75      PRELIMINARY;      PRT;   234 AA.
AC   Q8BN75;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Mus musculus 0 day neonate eyeball cDNA, RIKEN full-length enriched
DE   library, clone:E130020I02 product:activating transcription factor 2,
DE   full insert sequence.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Eyeball;
RX   MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA   Carninci P., Hayashizaki Y.;
RT   "High-efficiency full-length cDNA cloning.";
RL   Meth. Enzymol. 303:19-44(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Eyeball;
RX   MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA   RIKEN FANTOM Consortium;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Eyeball;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Eyeball;
RX   MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA   Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA   Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT   "Normalization and subtraction of cap-trapper-selected cDNAs to
RT   prepare full-length cDNA libraries for rapid discovery of new genes.";
RL   Genome Res. 10:1617-1630(2000).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Eyeball;
RX   MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA   Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA   Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA   Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA   Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA   Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA   Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA   Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT   "RIKEN integrated sequence analysis (RISA) system-384-format
RT   sequencing pipeline with 384 multicapillary sequencer.";
RL   Genome Res. 10:1757-1771(2000).
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Eyeball;
RA   Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA   Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA   Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA   Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA   Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA   Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA   Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA   Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA   Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA   Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA   Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK087424; BAC39869.1; -.
DR   HSSP; P15336; 1BHI.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0003676; F:nucleic acid binding; IEA.
DR   GO; GO:0008270; F:zinc ion binding; IEA.
DR   InterPro; IPR007087; Znf_C2H2.
DR   Pfam; PF00096; zf-C2H2; 1.
DR   SMART; SM00355; ZnF_C2H2; 1.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 1.
SQ   SEQUENCE   234 AA;  25208 MW;  F06750FA9EB37A4D CRC64;

  Query Match             90.8%;  Score 984;  DB 2;  Length 234;
  Best Local Similarity   100.0%;  Pred. No. 2.8e-64;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         19 MSDDKPFLCTAPGCGQRFTNEDHLAVHKHKHEMTLKFGPARNDSVIVADQTPTPTRFLKN 78
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSDDKPFLCTAPGCGQRFTNEDHLAVHKHKHEMTLKFGPARNDSVIVADQTPTPTRFLKN 60

Qy         79 CEEVGLFNELASPFENEFKKASEDDIKKMPLDLSPLATPIIRSKIEEPSVVETTHQDSPL 138
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CEEVGLFNELASPFENEFKKASEDDIKKMPLDLSPLATPIIRSKIEEPSVVETTHQDSPL 120

Qy        139 PHPESTTSDEKEVPLAQTAQPTSAIVRPASLQVPNVLLTSSDSSVIIQQAVPSPTSSTVI 198
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 PHPESTTSDEKEVPLAQTAQPTSAIVRPASLQVPNVLLTSSDSSVIIQQAVPSPTSSTVI 180

Qy        199 TQAPSSNRPIV 209
              |||||||||||
Db        181 TQAPSSNRPIV 191


RESULT 3
Q8CBR9
ID   Q8CBR9      PRELIMINARY;      PRT;   420 AA.
AC   Q8CBR9;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Mus musculus adult male urinary bladder cDNA, RIKEN full-length
DE   enriched library, clone:9530046I22 product:activating transcription
DE   factor 2, full insert sequence.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Urinary bladder;
RX   MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA   Carninci P., Hayashizaki Y.;
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 8, 2005, 08:05:04 ; Search time 3170 Seconds
(without alignments)
2509.600 Million cell updates/sec

Title: AAH26175
Perfect score: 1084
Sequence: 1 MKFKLHVNSARQYKDLWNMS.....PSPTSSTVITQAPSSNRPIV 209

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/AAH26175/runat_07072005_101604_18123/app_query.fasta_1.391
-DB=EST -QFMT=fastap -SUFFIX=std.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=AAH26175_@CGN_1_1_4352_@runat_07072005_101604_18123 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1060	97.8	1606	3	AK087424	AK087424 Mus muscu
2	1056	97.4	904	7	CR763073	CR763073 DKFZp469I
3	1052	97.0	956	5	BU839869	BU839869 AGENCOURT
4	1046.5	96.5	3623	3	AK035438	AK035438 Mus muscu
5	1016	93.7	883	4	BI854668	BI854668 603381603
6	1014	93.5	900	5	BU158697	BU158697 AGENCOURT
7	944	87.1	805	1	AJ454962	AJ454962 AJ454962
8	943.5	87.0	826	1	AJ454852	AJ454852 AJ454852
9	923	85.1	792	6	CA750476	CA750476 UI-M-FY0-
10	920	84.9	744	1	AJ453735	AJ453735 AJ453735
11	918	84.7	802	7	CR629564	CR629564 DKFZp469G
12	915.5	84.5	716	2	BF613172	BF613172 de07c03.y
13	915.5	84.5	893	5	BX851498	BX851498 BX851498
14	910	83.9	746	6	CD348287	CD348287 UI-M-FY0-
15	910	83.9	749	6	CD348288	CD348288 UI-M-FY0-
16	895	82.6	742	7	CK634629	CK634629 UI-M-HN0-
17	883	81.5	664	7	CF535749	CF535749 UI-M-GH0-
18	883	81.5	782	6	CD348295	CD348295 UI-M-FY0-
19	873.5	80.6	909	2	BF166515	BF166515 601776614
20	866	79.9	670	7	CN367891	CN367891 170004247
21	863	79.6	650	7	CN367893	CN367893 170005326
22	860	79.3	620	5	BP220000	BP220000 BP220000
23	859	79.2	686	2	BE882603	BE882603 601507056
24	855.5	78.9	915	5	BQ715489	BQ715489 AGENCOURT
25	848.5	78.3	781	6	CA512317	CA512317 UI-R-FJ0-
26	845	78.0	875	5	BQ891414	BQ891414 AGENCOURT
27	842	77.7	758	6	CB521135	CB521135 UI-M-GH0-
28	841	77.6	552	9	CL570065	CL570065 AS1007 Sa
29	840	77.5	757	7	CK635513	CK635513 UI-M-HN0-
30	837	77.2	613	5	BU662818	BU662818 cl91g01.z
31	831.5	76.7	2106	3	AK051591	AK051591 Mus muscu
32	826	76.2	908	5	BQ229140	BQ229140 AGENCOURT
33	810	74.7	806	6	CB961368	CB961368 AGENCOURT
34	793	73.2	2835	3	AY724488	AY724488 Rattus no
35	791.5	73.0	758	6	CA510974	CA510974 UI-R-FJ0-
36	791	73.0	924	4	BI551997	BI551997 603195125
37	785	72.4	818	7	CN455466	CN455466 UI-M-HN0-
38	779	71.9	545	2	BE549191	BE549191 601078730
39	779	71.9	577	7	CN367889	CN367889 170004241
40	779	71.9	651	7	CN367890	CN367890 170005319
41	779	71.9	841	5	BU598130	BU598130 AGENCOURT
42	773.5	71.4	958	5	BU503620	BU503620 AGENCOURT
43	770	71.0	606	5	BQ037883	BQ037883 pgn1c.pk0
44	749	69.1	827	4	BI548366	BI548366 603189652
45	744	68.6	722	7	CF532094	CF532094 UI-M-FY0-

ALIGNMENTS

RESULT 1
AK087424
LOCUS AK087424 1606 bp mRNA linear HTC 03-APR-2004
DEFINITION Mus musculus 0 day neonate eyeball cDNA, RIKEN full-length enriched library, clone:E130020I02 product:activating transcription factor 2, full insert sequence.
ACCESSION AK087424
VERSION AK087424.1 GI:26352477
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,

```
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5
  AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6  (bases 1 to 1606)
  AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
            Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
            Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
            Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
            Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
            Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
            Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
            Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
            Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,
            URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
            Please visit our web site for further details.
            URL:http://genome.gsc.riken.jp/
            URL:http://fantom.gsc.riken.jp/.
FEATURES             Location/Qualifiers
     source          1. .1606
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:E130020I02"
                     /db_xref="taxon:10090"
                     /clone="E130020I02"
                     /tissue_type="eyeball"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="0 day neonate"
     CDS             349. .1053
                     /note="unnamed protein product; activating transcription
                     factor 2 (MGD|MGI:109349, GB|S76655, evidence: BLASTN,
                     99%, match=1432)
                     putative"
                     /codon_start=1
                     /protein_id="BAC39869.1"
                     /db_xref="GI:26352478"
                     /translation="MSDDKPFLCTAPGCGQRFTNEDHLAVHKHKHEMTLKFGPARNDS
                     VIVADQTPTPTRFLKNCEEVGLFNELASPFENEFKKASEDDIKKMPLDLSPLATPIIR
                     SKIEEPSVVETTHQDSPLPHPESTTSDEKEVPLAQTAQPTSAIVRPASLQVPNVLLTS
                     SDSSVIIQQAVPSPTSSTVITQAPSSNRPIVPVPGPFPLLLHLPNGQTMPVAIPASIT
                     SSNVHVPAAVPVSWKV"
ORIGIN
```

```
Alignment Scores:
Pred. No.:              7.53e-99        Length:         1606
Score:                  1060.00         Matches:        206
Percent Similarity:     98.56%          Conservative:   0
Best Local Similarity:  98.56%          Mismatches:     3
Query Match:            97.79%          Indels:         0
DB:                     3               Gaps:           0

AAH26175 (1-209) x AK087424 (1-1606)

Qy          1 MetLysPheLysLeuHisValAsnSerAlaArgGlnTyrLysAspLeuTrpAsnMetSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        295 ATGAAATTCAAGTTGCATGTGAATTCTGCCAGATAAAACACGGACCTGTGGAATATGAGT 354

Qy         21 AspAspLysProPheLeuCysThrAlaProGlyCysGlyGlnArgPheThrAsnGluAsp 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        355 GATGACAAACCCTTTCTATGCACTGCCCCTGGGTGTGGCCAGCGTTTTACCAACGAGGAT 414

Qy         41 HisLeuAlaValHisLysHisLysHisGluMetThrLeuLysPheGlyProAlaArgAsn 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        415 CATTTGGCTGTCCATAAACATAAACATGAGATGACACTGAAATTTGGTCCAGCACGTAAT 474

Qy         61 AspSerValIleValAlaAspGlnThrProThrProThrArgPheLeuLysAsnCysGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        475 GACAGTGTCATTGTGGCTGATCAGACTCCAACGCCAACAAGATTCCTAAAAAACTGTGAA 534

Qy         81 GluValGlyLeuPheAsnGluLeuAlaSerProPheGluAsnGluPheLysLysAlaSer 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        535 GAAGTGGGTTTGTTCAATGAGTTGGCAAGTCCATTTGAAAATGAATTCAAGAAGGCTTCC 594

Qy        101 GluAspAspIleLysLysMetProLeuAspLeuSerProLeuAlaThrProIleIleArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        595 GAAGATGACATTAAAAAAATGCCTCTAGATTTGTCCCCTCTTGCAACACCCATCATAAGA 654

Qy        121 SerLysIleGluGluProSerValValGluThrThrHisGlnAspSerProLeuProHis 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        655 AGCAAAATTGAGGAGCCTTCTGTTGTAGAAACAACTCACCAGGACAGCCCTTTACCTCAC 714

Qy        141 ProGluSerThrThrSerAspGluLysGluValProLeuAlaGlnThrAlaGlnProThr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        715 CCCGAGTCGACTACCAGTGATGAAAAGGAAGTACCATTGGCACAAACTGCACAGCCCACA 774

Qy        161 SerAlaIleValArgProAlaSerLeuGlnValProAsnValLeuLeuThrSerSerAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        775 TCAGCTATCGTTCGTCCAGCATCATTACAGGTTCCCAATGTGCTGCTCACAAGTTCTGAC 834

Qy        181 SerSerValIleIleGlnGlnAlaValProSerProThrSerSerThrValIleThrGln 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        835 TCAAGTGTAATTATTCAACAAGCAGTACCTTCACCAACCTCAAGTACTGTAATCACCCAG 894

Qy        201 AlaProSerSerAsnArgProIleVal 209
              ||||||||||||||||||||||||||
Db        895 GCACCATCCTCTAACAGGCCAATTGTT 921

RESULT 2
CR763073
LOCUS       CR763073                904 bp    mRNA    linear   EST 23-SEP-2004
DEFINITION  DKFZp469I1824_r1 469 (synonym: pkid1) Pongo pygmaeus cDNA clone
            DKFZp469I1824 5', mRNA sequence.
ACCESSION   CR763073
VERSION     CR763073.1  GI:52600802
KEYWORDS    EST.
SOURCE      Pongo pygmaeus (orangutan)
  ORGANISM  Pongo pygmaeus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
REFERENCE   1  (bases 1 to 904)
  AUTHORS   Ansorge,W., Krieger,S., Regiert,T., Rittmueller,C., Schwager,B.,
            Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
            Wiemann,S.
  TITLE     Pongo pygmaeus mRNA (Ansorge,W., Krieger,S., Regiert,T., et al.)
  JOURNAL   Unpublished (2004)
COMMENT     Contact: MIPS
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 8, 2005, 03:11:06 ; Search time 28 Seconds
(without alignments)
718.189 Million cell updates/sec

Title: AAH26175
Perfect score: 1084
Sequence: 1 MKFKLHVNSARQYKDLWNMS.....PSPTSSTVITQAPSSNRPIV 209

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1084	100.0	505	1	S05380	transcription fact
2	923	85.1	486	2	JC4028	activating transcr
3	798	73.6	456	2	B42026	cyclic AMP respons
4	719	66.3	448	2	A42026	cAMP response elem
5	711	65.6	389	1	A39429	cAMP response elem
6	561	51.8	483	2	S12741	transcription fact
7	528	48.7	358	2	C42026	cyclic AMP respons
8	390.5	36.0	508	2	A45477	cAMP response elem
9	297	27.4	313	2	A34785	DNA-binding protei
10	133.5	12.3	593	2	JC7829	metal-responsive t
11	115	10.6	565	2	T39863	zinc finger protei
12	110.5	10.2	1367	1	S48478	glucan 1,4-alpha-g
13	110	10.1	447	2	I56511	Zic protein - mous
14	109.5	10.1	428	2	T43532	zinc finger protei
15	109.5	10.1	673	2	S35335	transcription fact
16	106	9.8	1251	2	S49645	probable membrane
17	104.5	9.6	700	1	S01991	transforming prote
18	104	9.6	1046	2	T29776	hypothetical prote
19	104	9.6	1983	2	T00385	KIAA0624 protein -
20	103	9.5	1133	2	S54496	probable membrane
21	102.5	9.5	533	2	JS0304	developmental cont
22	102.5	9.5	644	2	S39356	transcription fact
23	102.5	9.5	672	1	S73715	cytadherence acces
24	102	9.4	1361	2	T30884	neural specific DN
25	101	9.3	907	1	QQBE21	membrane antigen g
26	100	9.2	474	2	S66480	carbon catabolite
27	100	9.2	686	1	S28050	transforming prote
28	100	9.2	1263	2	T13805	spalt-related prot
29	100	9.2	1350	2	T30341	zinc finger protei
30	99	9.1	394	2	S50486	hypothetical prote
31	98.5	9.1	457	2	I54340	DNA-binding protei
32	98	9.0	462	2	T11637	hypothetical prote
33	97.5	9.0	515	2	A44256	zinc-finger protei
34	97	8.9	382	2	S71669	finger protein MIG
35	97	8.9	532	2	C87793	protein C27A12.3 [
36	97	8.9	878	2	T21621	hypothetical prote
37	97	8.9	3375	2	T19821	hypothetical prote
38	96.5	8.9	1355	2	S40022	spalt protein - fr
39	96	8.9	609	2	A49839	odd-paired - fruit
40	96	8.9	753	2	S48059	metal-regulatory t
41	96	8.9	2109	2	E89066	protein H05O09.1 [
42	96	8.9	2109	2	T33247	hypothetical prote
43	95	8.8	185	2	T49660	glucan 1,4-alpha-g
44	95	8.8	744	2	T20969	hypothetical prote
45	95	8.8	1323	2	T30253	spalt protein - mo

ALIGNMENTS

RESULT 1
S05380
transcription factor ATF2 - human
N;Alternate names: activating transcription factor 2 (ATF-2); cAMP response element-bin
N;Contains: cAMP response element-binding protein HB16
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: S05380; A34776; B34223
R;Maekawa, T.; Sakura, H.; Kanei-Ishii, C.; Sudo, T.; Yoshimura, T.; Fujisawa, J.I.; Yo
EMBO J. 8, 2023-2028, 1989
A;Title: Leucine zipper structure of the protein CRE-BP1 binding to the cyclic AMP resp
A;Reference number: S05380; MUID:90005408; PMID:2529117
A;Accession: S05380
A;Molecule type: mRNA
A;Residues: 1-505 <MAE>
A;Cross-references: UNIPROT:P15336; EMBL:X15875; NID:g30214; PIDN:CAA33886.1; PID:g3021
R;Kara, C.J.; Liou, H.C.; Ivashkiv, L.B.; Glimcher, L.H.
Mol. Cell. Biol. 10, 1347-1357, 1990
A;Title: A cDNA for a human cyclic AMP response element-binding protein which is distin
A;Reference number: A34776; MUID:90205810; PMID:2320002
A;Accession: A34776
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 211-222,'N',224-505 <KAR>
A;Cross-references: GB:M31630; NID:g183787; PIDN:AAA35951.1; PID:g386762
R;Hai, T.; Liu, F.; Coukos, W.J.; Green, M.R.
Genes Dev. 3, 2083-2090, 1989
A;Title: Transcription factor ATF cDNA clones: an extensive family of leucine zipper pr
A;Reference number: A91622; MUID:90185187; PMID:2516827
A;Accession: B34223
A;Status: nucleic acid sequence not shown; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 107-357,'V',359-465 <HA2>
C;Genetics:
A;Gene: GDB:ATF2; CREB2; TREB7; CRE-BP1
A;Cross-references: GDB:128011; OMIM:123811
A;Map position: 2q32-2q32
C;Superfamily: cAMP response element-binding protein 1; fos/jun DNA-binding domain homo
C;Keywords: alternative splicing; DNA binding; nucleus; transcription regulation
F;347-387/Domain: fos/jun DNA-binding domain homology <FJD>

Query Match 100.0%; Score 1084; DB 1; Length 505;
Best Local Similarity 100.0%; Pred. No. 5.3e-78;
Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MKFKLHVNSARQYKDLWNMSDDKPFLCTAPGCGQRFTNEDHLAVHKHKHEMTLKFGPARN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKFKLHVNSARQYKDLWNMSDDKPFLCTAPGCGQRFTNEDHLAVHKHKHEMTLKFGPARN 60

Qy         61 DSVIVADQTPTPTRFLKNCEEVGLFNELASPFENEFKKASEDDIKKMPLDLSPLATPIIR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DSVIVADQTPTPTRFLKNCEEVGLFNELASPFENEFKKASEDDIKKMPLDLSPLATPIIR 120
```

```
Qy        121 SKIEEPSVVETTHQDSPLPHPESTTSDEKEVPLAQTAQPTSAIVRPASLQVPNVLLTSSD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SKIEEPSVVETTHQDSPLPHPESTTSDEKEVPLAQTAQPTSAIVRPASLQVPNVLLTSSD 180

Qy        181 SSVIIQQAVPSPTSSTVITQAPSSNRPIV 209
              |||||||||||||||||||||||||||||
Db        181 SSVIIQQAVPSPTSSTVITQAPSSNRPIV 209


RESULT 2
JC4028
activating transcription factor 2 - African clawed frog
N;Alternate names: cyclic AMP-response element-binding protein
C;Species: Xenopus laevis (African clawed frog)
C;Date: 13-Jun-1995 #sequence_revision 14-Jul-1995 #text_change 09-Jul-2004
C;Accession: JC4028
R;Villarreal, X.C.; Richter, J.D.
Gene 153, 225-229, 1995
A;Title: Analysis of ATF2 gene expression during early Xenopus laevis developme nt.
A;Reference number: JC4028; MUID:95180723; PMID:7875593
A;Accession: JC4028
A;Molecule type: mRNA
A;Residues: 1-486 <VIL>
A;Cross-references: UNIPROT:Q91576; GB:U16158; NID:g887779; PIDN:AAA69518.1; PID:g710326
C;Comment: This protein is a sequence-specific DNA-binding protein that mediates transcr
C;Superfamily: cAMP response element-binding protein 1; fos/jun DNA-binding domain homol
C;Keywords: leucine zipper; phosphoprotein; transcription regulation
F;37-398/Region: leucine zipper motif
F;329-369/Domain: fos/jun DNA-binding domain homology <FJD>
F;332-364/Region: basic
F;82/Binding site: phosphate (Ser) (covalent) (by protein kinase A) #status predicted
F;102/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted

  Query Match             85.1%;  Score 923;  DB 2;  Length 486;
  Best Local Similarity   94.8%;  Pred. No. 2.6e-65;
  Matches  182;  Conservative    5;  Mismatches    3;  Indels    2;  Gaps    2;

Qy         19 MSDDKPFLCTAPGCGQRFTNEDHLAVHKHKHEMTLKFGPARNDSVIVADQTPTPTRFLKN 78
              ||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db          1 MSDDKPFLCTAPGCGQRFTNKDHLAVHKHKHEMTLKFGPARNDSVIVADQTPTPTRFLKN 60

Qy         79 CEEVGLFNELASPFENEFKKASEDDIKKMPLDLSPLATPIIRSKIEEPSVVETTHQDSPL 138
              |||||||||||||||||||||:| ||||||||||||||||||||||||||||||||||||
Db         61 CEEVGLFNELASPFENEFKKASEED-KKMPLDLSPLATPIIRSKIEEPSVVETTHQDSPL 119

Qy        139 PHPESTTSDE-KEVPLAQTAQPTSAIVRPASLQVPNVLLTSSDSSVIIQQAVPSPTSSTV 197
              |||||||||:  :||||||||||||||||||||||||||||||||||||||||||||||
Db        120 PHPESTTSDQHNQVPLAQTAQPTSAIVRPASLQVPNVLLTSSDSSVIIQQAVPSPTSSTV 179

Qy        198 ITQAPSSNRPIV 209
              ||||||:|  ||
Db        180 ITQAPSTNSSIV 191


RESULT 3
B42026
cyclic AMP response element DNA-binding protein isoform 1 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 26-Aug-1999
C;Accession: B42026
R;Georgopoulos, K.; Morgan, B.A.; Moore, D.D.
Mol. Cell. Biol. 12, 747-757, 1992
A;Title: Functionally distinct isoforms of the CRE-BP DNA-binding protein mediate activi
A;Reference number: A42026; MUID:92123199; PMID:1531087
A;Contents: EL4
A;Accession: B42026
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-456 <GEO>
A;Cross-references: GB:S76657; NID:g243428; PIDN:AAB21128.1; PID:g243429
A;Note: sequence extracted from NCBI backbone (NCBIN:76657, NCBIP:76658)
C;Superfamily: cAMP response element-binding protein 1; fos/jun DNA-binding domain homol
F;297-337/Domain: fos/jun DNA-binding domain homology <FJD>

  Query Match             73.6%;  Score 798;  DB 2;  Length 456;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-55;
  Matches  159;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         51 MTLKFGPARNDSVIVADQTPTPTRFLKNCEEVGLFNELASPFENEFKKASEDDIKKMPLD 110
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MTLKFGPARNDSVIVADQTPTPTRFLKNCEEVGLFNELASPFENEFKKASEDDIKKMPLD 60

Qy        111 LSPLATPIIRSKIEEPSVVETTHQDSPLPHPESTTSDEKEVPLAQTAQPTSAIVRPASLQ 170
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LSPLATPIIRSKIEEPSVVETTHQDSPLPHPESTTSDEKEVPLAQTAQPTSAIVRPASLQ 120

Qy        171 VPNVLLTSSDSSVIIQQAVPSPTSSTVITQAPSSNRPIV 209
              |||||||||||||||||||||||||||||||||||||||
Db        121 VPNVLLTSSDSSVIIQQAVPSPTSSTVITQAPSSNRPIV 159


RESULT 4
A42026
cAMP response element-binding protein 3 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 26-Aug-1999
C;Accession: A42026
R;Georgopoulos, K.; Morgan, B.A.; Moore, D.D.
Mol. Cell. Biol. 12, 747-757, 1992
A;Title: Functionally distinct isoforms of the CRE-BP DNA-binding protein mediate activi
A;Reference number: A42026; MUID:92123199; PMID:1531087
A;Contents: EL4
A;Accession: A42026
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-448 <GEO>
A;Cross-references: GB:S76655; NID:g243426; PIDN:AAB21127.1; PID:g243427
A;Note: sequence extracted from NCBI backbone (NCBIN:76655, NCBIP:76656)
C;Superfamily: cAMP response element-binding protein 1; fos/jun DNA-binding domain homol
F;289-329/Domain: fos/jun DNA-binding domain homology <FJD>

  Query Match             66.3%;  Score 719;  DB 2;  Length 448;
  Best Local Similarity   100.0%;  Pred. No. 2.9e-49;
  Matches  143;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         67 DQTPTPTRFLKNCEEVGLFNELASPFENEFKKASEDDIKKMPLDLSPLATPIIRSKIEEP 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          9 DQTPTPTRFLKNCEEVGLFNELASPFENEFKKASEDDIKKMPLDLSPLATPIIRSKIEEP 68

Qy        127 SVVETTHQDSPLPHPESTTSDEKEVPLAQTAQPTSAIVRPASLQVPNVLLTSSDSSVIIQ 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         69 SVVETTHQDSPLPHPESTTSDEKEVPLAQTAQPTSAIVRPASLQVPNVLLTSSDSSVIIQ 128

Qy        187 QAVPSPTSSTVITQAPSSNRPIV 209
              |||||||||||||||||||||||
Db        129 QAVPSPTSSTVITQAPSSNRPIV 151


RESULT 5
A39429
cAMP response element-binding protein ATF2 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: A39429
R;Kageyama, R.; Sasai, Y.; Nakanishi, S.
J. Biol. Chem. 266, 15525-15531, 1991
A;Title: Molecular characterization of transcription factors that bind to the cAMP respo
A;Reference number: A39429; MUID:91332085; PMID:1714459
A;Accession: A39429
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-389 <KAG>
A;Cross-references: UNIPROT:Q00969; GB:M65148; NID:g206569; PIDN:AAA42013.1; PID:g206570
C;Superfamily: cAMP response element-binding protein 1; fos/jun DNA-binding domain homol
C;Keywords: DNA binding; nucleus; transcription regulation
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 8, 2005, 10:23:27 ; Search time 593 Seconds
(without alignments)
2212.762 Million cell updates/sec

Title: AAH26175
Perfect score: 1084
Sequence: 1 MKFKLHVNSARQYKDLWNMS.....PSPTSSTVITQAPSSNRPIV 209

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 6330943 seqs, 3139157217 residues

Total number of hits satisfying chosen parameters: 12661886

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/AAH26175/runat_07072005_101607_18236/app_query.fasta_1.391
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=std.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=AAH26175_@CGN_1_1_741_@runat_07072005_101607_18236
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	569	52.5	2755	20	US-10-357-930-22633	Sequence 22633, A
2	569	52.5	2755	20	US-10-357-930-28476	Sequence 28476, A
3	434	40.0	800	20	US-10-357-930-23141	Sequence 23141, A
4	434	40.0	800	20	US-10-357-930-29011	Sequence 29011, A
5	427	39.4	501	9	US-09-833-790-215	Sequence 215, App
6	390.5	36.0	2637	11	US-09-968-007A-974	Sequence 974, App
7	390.5	36.0	2637	15	US-10-171-581-285	Sequence 285, App
8	390.5	36.0	2637	21	US-10-843-641A-7444	Sequence 7444, Ap
9	373.5	34.5	2345	20	US-10-723-860-7409	Sequence 7409, Ap
10	332.5	30.7	432	20	US-10-357-930-35258	Sequence 35258, A
11	332.5	30.7	432	20	US-10-357-930-44098	Sequence 44098, A
12	332	30.6	397	20	US-10-357-930-14137	Sequence 14137, A
13	210.5	19.4	509	16	US-10-210-120-82	Sequence 82, Appl
14	210.5	19.4	509	22	US-10-909-035-82	Sequence 82, Appl
15	180.5	16.7	273	20	US-10-357-930-4968	Sequence 4968, Ap
16	116.5	10.7	4168	17	US-10-108-260A-1212	Sequence 1212, Ap
17	116.5	10.7	6682	20	US-10-723-860-5563	Sequence 5563, Ap
18	115.5	10.7	2493	20	US-10-425-115-114107	Sequence 114107,
19	115	10.6	2193	20	US-10-425-115-114963	Sequence 114963,
20	112.5	10.4	3511	20	US-10-723-860-4919	Sequence 4919, Ap
21	112	10.3	3138	9	US-09-880-107-1716	Sequence 1716, Ap
22	112	10.3	3138	17	US-10-172-118-909	Sequence 909, App
23	112	10.3	3138	18	US-10-342-887-909	Sequence 909, App
24	112	10.3	3138	20	US-10-723-860-176	Sequence 176, App
25	111.5	10.3	2976	20	US-10-425-115-47091	Sequence 47091, A
26	110.5	10.2	4104	9	US-09-801-368-107	Sequence 107, App
27	110.5	10.2	4104	19	US-10-793-639-44	Sequence 44, Appl
28	110.5	10.2	5262	16	US-10-240-965-79	Sequence 79, Appl
29	108.5	10.0	1448	9	US-09-864-761-21928	Sequence 21928, A
30	108.5	10.0	1780	17	US-10-108-260A-874	Sequence 874, App
31	108.5	10.0	1947	9	US-09-864-761-4439	Sequence 4439, Ap
32	108.5	10.0	4067	19	US-10-250-889-27	Sequence 27, Appl
33	108	10.0	2979	20	US-10-425-115-113191	Sequence 113191,
34	107.5	9.9	2487	20	US-10-425-115-112880	Sequence 112880,
35	107.5	9.9	5637	20	US-10-425-115-115805	Sequence 115805,
36	107.5	9.9	18717	15	US-10-156-761-927	Sequence 927, App
37	107.5	9.9	30690	19	US-10-204-862A-1	Sequence 1, Appli
38	107.5	9.9	9025608	15	US-10-156-761-1	Sequence 1, Appli
39	107	9.9	2457	17	US-10-305-720-1106	Sequence 1106, Ap
40	106	9.8	1825	15	US-10-244-367-41	Sequence 41, Appl
41	106	9.8	2106	10	US-09-918-624B-57	Sequence 57, Appl
C 42	105.5	9.7	842	16	US-10-029-386-22926	Sequence 22926, A
43	105.5	9.7	2334	20	US-10-425-115-113926	Sequence 113926,
44	105.5	9.7	2658	20	US-10-425-115-47126	Sequence 47126, A
45	105.5	9.7	3744	19	US-10-115-635-241	Sequence 241, App

ALIGNMENTS

RESULT 1
US-10-357-930-22633
; Sequence 22633, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel. Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16

```
;  PRIOR APPLICATION NUMBER: 60/183,319
;  PRIOR FILING DATE: 2000-02-17
;  PRIOR APPLICATION NUMBER: 60/189,862
;  PRIOR FILING DATE: 2000-03-16
;  PRIOR APPLICATION NUMBER: 60/207,454
;  PRIOR FILING DATE: 2000-05-25
;  PRIOR APPLICATION NUMBER: 60/211,314
;  PRIOR FILING DATE: 2000-06-09
;  PRIOR APPLICATION NUMBER: 60/219,007
;  PRIOR FILING DATE: 2000-07-18
;  PRIOR APPLICATION NUMBER: 60/255,281
;  PRIOR FILING DATE: 2000-12-13
;  NUMBER OF SEQ ID NOS: 62232
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 22633
;   LENGTH: 2755
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 2753, 2754, 2755
;   OTHER INFORMATION: n = A,T,C or G
US-10-357-930-22633

Alignment Scores:
Pred. No.:              3.02e-57      Length:          2755
Score:                  569.00        Matches:         126
Percent Similarity:     69.90%        Conservative:    18
Best Local Similarity:  61.17%        Mismatches:      42
Query Match:            52.49%        Indels:          20
DB:                     20            Gaps:            7

AAH26175 (1-209) x US-10-357-930-22633 (1-2755)

Qy         11 ArgGlnTyrLysAspLeuTrp-----------AsnMetSerAspAspLysProPheLeu 26
              |||   |||:::  |||             ||||||  ||||||:::||||||:::
Db         65 CGGAGCTACCAGCCGCTCCTCTCTGCTATATGAAATATGGGAGACGACAGACCGTTTGTG 124

Qy         27 CysThrAlaProGlyCysGlyGlnArgPheThrAsnGluAspHisLeuAlaValHisLys 46
              |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 TGCAATGCCCCGGGCTGTGGACAGAGATTTACAAACGAGGACCACCTGGCAGTTCATAAA 184

Qy         47 HisLysHisGluMetThrLeuLysPheGlyProAlaArgAsnAspSerValIleValAla 66
              ||||||||||||||||||||||||||||||||||||  ||||||||||||:::|||
Db        185 CACAAGCATGAGATGACATTGAAATTTGGCCCAGCCCGAACTGACTCAGTCATCATTGCA 244

Qy         67 AspGlnThrProThrProThrArgPheLeuLysAsnCysGluGluValGlyLeuPheAsn 86
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 GATCAAACGCCTACTCCAACTAGATTCCTGAAGAACTGTGAGGAGGTGGGACTCTTCAAT 304

Qy         87 GluLeuAlaSerProPheGluAsnGluPheLysLysAlaSerGluAspAspIleLysLys 106
              ||||||||||||  ||||||:::||||||||||||||:::::::::|||  ||||||
Db        305 GAACTAGCTAGCTCCTTTGAACATGAATTCAAGAAAGCTGCAGATGAGGATGAGAAAAAG 364

Qy        107 Met---------ProLeuAspLeuSerProLeuAlaThrProIleIleArgSerLysIle 123
                        ||||||||:::|||     :::||||||  |||:::  |||
Db        365 GCTGCTGCTGGGCCCCTTGACATGTCTCTGCCTTCCACACCAGACATCAAAATCAAAGAA 424

Qy        124 GluGluProSerValValGluThrThrHisGlnAspSerProLeuProHisProGluSer 143
              |||||||||      |||:::::::::      ||||||||          |||  |||
Db        425 GAAGAGCCAGTGGAGGTAGACTCATCCCCACCTGATAGCCCTGCCTCTAGTCCCTGTTCC 484

Qy        144 ThrThrSerAspGluLysGluVal---------ProLeuAlaGlnThrAlaGlnProThr 160
                       ||||||||||||         |||:::         :::  ||||||
Db        485 CCACCACTGAAGGAGAAGGAGGTTACCCCAAAGCCTGTTCTG---ATCTCTACCCCCACA 541

Qy        161 SerAlaIleValArgProAlaSerLeuGlnValProAsnValLeuLeuThrSerSerAsp 180
              ||||||||||||  ||||||                |||                |||
Db        542 CCCACCATTGTACGTCCTGGCTCCCTGCCT-----------CTCCACTTGGGCTATGAT 589

Qy        181 SerSerValIleIleGlnGlnAlaValProSerProThrSerSerThrValIleThrGln 200
                     :::         :::|||||||||||||||      ||||||||||||
Db        590 CCA---------CTTCATCCAACCCTTCCCTCCCCAACCTCT------GTCATCACACAG 634
```

```
Qy        201 AlaProSerSerAsnArg 206
              ||||||  |||||||||
Db        635 GCTCCACCATCCAACAGG 652

RESULT 2
US-10-357-930-28476
; Sequence 28476, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
;  APPLICANT: Schlegel. Robert
;  APPLICANT:  Endege, Wilson
;  APPLICANT:  Monahan, John
;  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
;  TITLE OF INVENTION:  IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
;  TITLE OF INVENTION:  HUMAN PROSTATE CANCER
;  FILE REFERENCE: MRI-007BCN
;  CURRENT APPLICATION NUMBER: US/10/357,930
;  CURRENT FILING DATE:  2003-02-04
;  PRIOR APPLICATION NUMBER: 09/785,276
;  PRIOR FILING DATE: 2003-02-16
;  PRIOR APPLICATION NUMBER: 60/183,319
;  PRIOR FILING DATE: 2000-02-17
;  PRIOR APPLICATION NUMBER: 60/189,862
;  PRIOR FILING DATE: 2000-03-16
;  PRIOR APPLICATION NUMBER: 60/207,454
;  PRIOR FILING DATE: 2000-05-25
;  PRIOR APPLICATION NUMBER: 60/211,314
;  PRIOR FILING DATE: 2000-06-09
;  PRIOR APPLICATION NUMBER: 60/219,007
;  PRIOR FILING DATE: 2000-07-18
;  PRIOR APPLICATION NUMBER: 60/255,281
;  PRIOR FILING DATE: 2000-12-13
;  NUMBER OF SEQ ID NOS: 62232
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 28476
;   LENGTH: 2755
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 2753, 2754, 2755
;   OTHER INFORMATION: n = A,T,C or G
US-10-357-930-28476

Alignment Scores:
Pred. No.:              3.02e-57      Length:          2755
Score:                  569.00        Matches:         126
Percent Similarity:     69.90%        Conservative:    18
Best Local Similarity:  61.17%        Mismatches:      42
Query Match:            52.49%        Indels:          20
DB:                     20            Gaps:            7

AAH26175 (1-209) x US-10-357-930-28476 (1-2755)

Qy         11 ArgGlnTyrLysAspLeuTrp-----------AsnMetSerAspAspLysProPheLeu 26
              |||   |||:::  |||             ||||||  ||||||:::||||||:::
Db         65 CGGAGCTACCAGCCGCTCCTCTCTGCTATATGAAATATGGGAGACGACAGACCGTTTGTG 124

Qy         27 CysThrAlaProGlyCysGlyGlnArgPheThrAsnGluAspHisLeuAlaValHisLys 46
              |||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 TGCAATGCCCCGGGCTGTGGACAGAGATTTACAAACGAGGACCACCTGGCAGTTCATAAA 184

Qy         47 HisLysHisGluMetThrLeuLysPheGlyProAlaArgAsnAspSerValIleValAla 66
              ||||||||||||||||||||||||||||||||||||  ||||||||||||:::|||
Db        185 CACAAGCATGAGATGACATTGAAATTTGGCCCAGCCCGAACTGACTCAGTCATCATTGCA 244

Qy         67 AspGlnThrProThrProThrArgPheLeuLysAsnCysGluGluValGlyLeuPheAsn 86
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 GATCAAACGCCTACTCCAACTAGATTCCTGAAGAACTGTGAGGAGGTGGGACTCTTCAAT 304

Qy         87 GluLeuAlaSerProPheGluAsnGluPheLysLysAlaSerGluAspAspIleLysLys 106
              ||||||||||||  ||||||:::||||||||||||||:::::::::|||  ||||||
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 8, 2005, 08:11:07 ; Search time 183 Seconds
(without alignments)
1868.753 Million cell updates/sec

Title: AAH26175
Perfect score: 1084
Sequence: 1 MKFKLHVNSARQYKDLWNMS.....PSPTSSTVITQAPSSNRPIV 209

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/AAH26175/runat_07072005_101605_18139/app_query.fasta_1.391
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=std.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=AAH26175_@CGN_1_1_177_@runat_07072005_101605_18139 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1084	100.0	1621	4	US-09-949-016-5410	Sequence 5410, Ap
2	407.5	37.6	65966	4	US-09-949-016-17152	Sequence 17152, A
3	263	24.3	229	4	US-09-513-999C-2946	Sequence 2946, Ap
c 4	148	13.7	601	4	US-09-949-016-189073	Sequence 189073,
c 5	143	13.2	601	4	US-09-949-016-189064	Sequence 189064,
6	112	10.3	3138	3	US-09-234-332-5	Sequence 5, Appli
7	110.5	10.2	4104	4	US-09-614-221A-44	Sequence 44, Appl
8	108.5	10.0	2394	4	US-09-949-016-5520	Sequence 5520, Ap
9	108.5	10.0	24428	4	US-09-949-016-17262	Sequence 17262, A
10	107	9.9	2457	4	US-09-016-434-1106	Sequence 1106, Ap
11	106	9.8	1825	4	US-09-342-325C-41	Sequence 41, Appl
12	104.5	9.6	2627	4	US-09-023-655-1251	Sequence 1251, Ap
13	104.5	9.6	2638	1	US-08-306-691B-46	Sequence 46, Appl
14	104.5	9.6	2653	4	US-09-976-594-66	Sequence 66, Appl
15	104	9.6	6745	4	US-09-774-528-286	Sequence 286, App
16	102.5	9.5	3068	2	US-08-224-482-1	Sequence 1, Appli
17	102.5	9.5	3086	1	US-08-040-548-15	Sequence 15, Appl
18	102.5	9.5	3086	1	US-08-466-344-15	Sequence 15, Appl
19	102.5	9.5	3086	6	5206152-1	Patent No. 5206152
20	102.5	9.5	3086	6	5206152-1	Patent No. 5206152
21	102	9.4	589	4	US-09-229-037-29	Sequence 29, Appl
22	102	9.4	589	4	US-09-478-681-29	Sequence 29, Appl
23	102	9.4	589	4	US-09-779-233-17	Sequence 17, Appl
24	102	9.4	589	4	US-09-706-243A-29	Sequence 29, Appl
25	102	9.4	2447	4	US-09-949-016-5274	Sequence 5274, Ap
26	101.5	9.4	3240	4	US-09-949-016-5548	Sequence 5548, Ap
27	101.5	9.4	11929	4	US-09-949-016-17290	Sequence 17290, A
28	101	9.3	3833	1	US-08-917-320-18	Sequence 18, Appl
29	101	9.3	3833	5	PCT-US95-04611A-18	Sequence 18, Appl
30	101	9.3	5931	3	US-08-783-774-1	Sequence 1, Appli
31	101	9.3	5931	3	US-09-556-706B-1	Sequence 1, Appli
32	101	9.3	5931	4	US-09-724-418A-1	Sequence 1, Appli
33	100.5	9.3	2901	4	US-09-342-325C-43	Sequence 43, Appl
34	100	9.2	292	4	US-09-395-448-14	Sequence 14, Appl
35	100	9.2	292	4	US-09-925-796-14	Sequence 14, Appl
36	100	9.2	292	4	US-09-941-450-14	Sequence 14, Appl
37	100	9.2	4674	4	US-09-410-551B-26	Sequence 26, Appl
38	100	9.2	4674	4	US-09-940-316B-26	Sequence 26, Appl
39	99.5	9.2	918	4	US-10-029-180-59	Sequence 59, Appl
40	99.5	9.2	136058	4	US-09-949-016-12565	Sequence 12565, A
41	99.5	9.2	136480	4	US-09-949-016-17064	Sequence 17064, A
42	99	9.1	7785	4	US-09-573-080A-32	Sequence 32, Appl
43	99	9.1	24735	4	US-09-949-016-12850	Sequence 12850, A
44	98.5	9.1	2655	4	US-09-799-451-187	Sequence 187, App
45	97	8.9	3020	4	US-09-220-132-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1
US-09-949-016-5410
; Sequence 5410, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5410
; LENGTH: 1621
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-5410

Alignment Scores:

Pred. No.:	8.96e-121	Length:	1621
Score:	1084.00	Matches:	209
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

AAH26175 (1-209) x US-09-949-016-5410 (1-1621)

Qy 1 MetLysPheLysLeuHisValAsnSerAlaArgGlnTyrLysAspLeuTrpAsnMetSer 20

```
Db         24 ATGAAATTCAAGTTACATGTGAATTCTGCCAGGCAATACAAGGACCTGTGGAATATGAGT 83

Qy         21 AspAspLysProPheLeuCysThrAlaProGlyCysGlyGlnArgPheThrAsnGluAsp 40

Db         84 GATGACAAACCCTTTCTATGTACTGCGCCTGGATGTGGCCAGCGTTTTACCAACGAGGAT 143

Qy         41 HisLeuAlaValHisLysHisLysHisGluMetThrLeuLysPheGlyProAlaArgAsn 60

Db        144 CATTTGGCTGTCCATAAACATAAACATGAGATGACACTGAAATTTGGTCCAGCACGTAAT 203

Qy         61 AspSerValIleValAlaAspGlnThrProThrProThrArgPheLeuLysAsnCysGlu 80

Db        204 GACAGTGTCATTGTGGCTGATCAGACCCCAACACCAACAAGATTCTTGAAAAACTGTGAA 263

Qy         81 GluValGlyLeuPheAsnGluLeuAlaSerProPheGluAsnGluPheLysLysAlaSer 100

Db        264 GAAGTGGGTTTGTTTAATGAGTTGGCGAGTCCATTTGAGAATGAATTCAAGAAAGCTTCA 323

Qy        101 GluAspAspIleLysLysMetProLeuAspLeuSerProLeuAlaThrProIleIleArg 120

Db        324 GAAGATGACATTAAAAAAATGCCTCTAGATTTATCCCCTCTTGCAACACCTATCATAAGA 383

Qy        121 SerLysIleGluGluProSerValValGluThrThrHisGlnAspSerProLeuProHis 140

Db        384 AGCAAAATTGAGGAGCCTTCTGTTGTAGAAACAACTCACCAGGATAGTCCTTTACCTCAC 443

Qy        141 ProGluSerThrThrSerAspGluLysGluValProLeuAlaGlnThrAlaGlnProThr 160

Db        444 CCAGAGTCTACTACCAGTGATGAGAAGGAAGTACCATTGGCACAAACTGCACAGCCCACA 503

Qy        161 SerAlaIleValArgProAlaSerLeuGlnValProAsnValLeuLeuThrSerSerAsp 180

Db        504 TCAGCTATTGTTCGTCCAGCATCATTACAGGTTCCCAATGTGCTGCTTACAAGTTCTGAC 563

Qy        181 SerSerValIleIleGlnGlnAlaValProSerProThrSerSerThrValIleThrGln 200

Db        564 TCAAGTGTAATTATTCAGCAGGCAGTACCTTCACCAACCTCAAGTACTGTAATCACCCAG 623

Qy        201 AlaProSerSerAsnArgProIleVal 209

Db        624 GCACCATCCTCTAACAGGCCAATTGTC 650
```

RESULT 2
US-09-949-016-17152
; Sequence 17152, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17152
; LENGTH: 65966
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17152

Alignment Scores:

Pred. No.:	1.89e-36	Length:	65966
Score:	407.50	Matches:	86
Percent Similarity:	61.38%	Conservative:	3
Best Local Similarity:	59.31%	Mismatches:	3
Query Match:	37.59%	Indels:	53
DB:	4	Gaps:	2

AAH26175 (1-209) x US-09-949-016-17152 (1-65966)

```
Qy         49 HisGluMetThrLeuLysPheGlyProAlaArgAsnAspSerValIleValAlaAspGln 68

Db      20072 CATGATTACTTCTTGACATTT-------------------------ATTTCAGATCAG 20104

Qy         69 ThrProThrProThrArgPheLeuLysAsnCysGluGluValGlyLeuPheAsnGluLeu 88

Db      20105 ACCCCAACACCAACAAGATTCTTGAAAAACTGTGAAGAAGTGGGTTTGTTTAATGAGTTG 20164

Qy         89 AlaSerProPheGluAsnGluPheLysLysAlaSerGluAspAspIleLysLys------ 106

Db      20165 GCGAGTCCATTTGAGAATGAATTCAAGAAAGCTTCAGAAGATGACATTAAAAAAGTATGT 20224

Qy        106 ---------------------------------------------------------- 106

Db      20225 TGTAACCATGAAATTTAAAAAGAGTACTTCTGACAGAAAAGTAATCATAGTTTATACTTC 20284

Qy        106 ---------------------------------------------------------- 106

Db      20285 CTGCTTTTTAAATTTAGACACACTGTTTTTGAACTTATAACCTTTTCCTTATTATTTGAA 20344

Qy        107 ------MetProLeuAspLeuSerProLeuAlaThrProIleIleArgSerLysIleGlu 124

Db      20345 TTATAGATGCCTCTAGATTTATCCCCTCTTGCAACACCTATCATAAGAAGCAAAATTGAG 20404

Qy        125 GluProSerValValGluThrThrHisGlnAspSerProLeuProHisProGluSerThr 144

Db      20405 GAGCCTTCTGTTGTAGAAACAACTCACCAGGATAGTCCTTTACCTCACCCAGAGTCTACT 20464

Qy        145 ThrSerAspGluLys 149

Db      20465 ACCAGTGATGAGAAG 20479
```

RESULT 3
US-09-513-999C-2946
; Sequence 2946, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 2946
; LENGTH: 229
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 72..227
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 87
; OTHER INFORMATION: y=c or t
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 6
; OTHER INFORMATION: Xaa=Pro or Ser
US-09-513-999C-2946

Alignment Scores:

Pred. No.:	6.4e-23	Length:	229

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 8, 2005, 03:47:39 ; Search time 530 Seconds
(without alignments)
2334.389 Million cell updates/sec

Title: AAH26175
Perfect score: 1084
Sequence: 1 MKFKLHVNSARQYKDLWNMS.....PSPTSSTVITQAPSSNRPIV 209

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/AAH26175/runat_07072005_101602_18107/app_query.fasta_1.391
-DB=N_Geneseq_16Dec04 -QFMT=fastap -SUFFIX=std.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=AAH26175_@CGN_1_1_644_@runat_07072005_101602_18107 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1084	100.0	1647	10	ADK67038	Adk67038 Gene #128
2	1084	100.0	1647	10	ADK66978	Adk66978 Gene #68
3	1062	98.0	1852	10	ADB53173	Adb53173 Primary r
4	940.5	86.8	3507	13	ADR08447	Adr08447 Full leng
5	571	52.7	1239	12	ADP80924	Adp80924 Mouse C13
6	569	52.5	2755	5	ABV28463	Abv28463 Human pro
7	569	52.5	2755	5	ABV22640	Abv22640 Human pro
8	566	52.2	3217	6	ABI99202	Abi99202 Mouse isc
9	434	40.0	800	5	ABV28993	Abv28993 Human pro
10	434	40.0	800	5	ABV23152	Abv23152 Human pro
11	427	39.4	501	6	AAS61674	Aas61674 Lung smal
12	390.5	36.0	2319	10	ADC18732	Adc18732 DNA of hu
13	390.5	36.0	2637	6	ABL69107	Abl69107 Kidney ca
14	390.5	36.0	2637	6	ABK83799	Abk83799 Human cDN
15	390.5	36.0	2637	10	ADH29017	Adh29017 Human chr
16	373.5	34.5	2345	12	ADQ24589	Adq24589 Human sof
17	332.5	30.7	432	5	ABV44079	Abv44079 Human pro
18	332.5	30.7	432	5	ABV35240	Abv35240 Human pro
19	332	30.6	397	5	ABV14146	Abv14146 Human pro
20	263	24.3	229	3	AAC02948	Aac02948 Human sec
21	210.5	19.4	509	10	ADD18510	Add18510 Human pro
22	180.5	16.7	273	5	ABV04977	Abv04977 Human pro
23	174	16.1	1080	4	ABL21331	Abl21331 Drosophil
24	170.5	15.7	3199	4	ABL21330	Abl21330 Drosophil
25	116.5	10.7	4168	11	ADM02527	Adm02527 Human cDN
26	116.5	10.7	6682	12	ADQ22743	Adq22743 Human sof
27	112.5	10.4	3511	12	ADQ22099	Adq22099 Human sof
28	112	10.3	3138	3	AAA62683	Aaa62683 Human Zic
29	112	10.3	3138	6	ABN95218	Abn95218 Gene #171
30	112	10.3	3138	12	ADN04235	Adn04235 Antipsori
31	112	10.3	3138	12	ADQ17359	Adq17359 Human sof
32	112	10.3	3138	13	ADR25048	Adr25048 Breast ca
33	111.5	10.3	2947	6	ABI99239	Abi99239 Mouse isc
34	110.5	10.2	4434	5	AAS79667	Aas79667 DNA encod
35	110.5	10.2	5262	6	AAS94824	Aas94824 Human DNA
36	109.5	10.1	12381	3	AAZ58381	Aaz58381 Streptomy
37	109	10.1	2429	4	ABL06615	Abl06615 Drosophil
c 38	109	10.1	9571	4	ABL06614	Abl06614 Drosophil
39	108.5	10.0	1448	4	AAI24539	Aai24539 Probe #14
40	108.5	10.0	1448	4	ABA69703	Aba69703 Human foe
41	108.5	10.0	1448	4	AAI49785	Aai49785 Probe #18
42	108.5	10.0	1448	4	ABA51616	Aba51616 Human bre
43	108.5	10.0	1448	4	ABA36608	Aba36608 Probe #15
44	108.5	10.0	1448	4	AAK43762	Aak43762 Human bon
45	108.5	10.0	1448	4	AAK17891	Aak17891 Human bra

ALIGNMENTS

RESULT 1
ADK67038
ID ADK67038 standard; DNA; 1647 BP.
XX
AC ADK67038;
XX
DT 06-MAY-2004 (first entry)
XX
DE Gene #128 for inhibitory RNA to manipulate stem cell phenotype.
XX
KW ds; gene; inhibitory RNA; RNAi; Notch signaling target gene; phenotype;
KW pluripotent stem cell.
XX
OS Homo sapiens.
XX
PN WO2003068961-A2.
XX
PD 21-AUG-2003.
XX
PF 12-FEB-2003; 2003WO-GB000579.
XX
PR 13-FEB-2002; 2002GB-00003359.
PR 13-FEB-2002; 2002GB-00003387.
XX
PA (AXOR-) AXORDIA LTD.
XX
PI Andrews P, Walsh J, Gokhale P;
XX
DR WPI; 2003-697528/66.

XX
PT New inhibitory RNA molecule having double stranded RNA molecules, useful
PT for manipulating the phenotype of stem cells, preferably pluripotential
PT stem cells.
XX
PS Disclosure; SEQ ID NO 128; 157pp; English.
XX
CC The invention relates to an inhibitory RNA (RNAi) molecule derived from a
CC nucleic acid molecule comprising a defined nucleic acid sequences given
CC in the specification or a sequence which hybridizes to the sequences and
CC encodes a Notch signaling target gene or which is a degenerate as a
CC result of the genetic code of the sequences. The methods and compositions
CC of the present invention are useful for manipulating the phenotype of
CC stem cells, preferably pluripotent stem cells. This sequence corresponds
CC to one of the nucleic acid molecules of the invention.
XX
SQ Sequence 1647 BP; 504 A; 396 C; 341 G; 406 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	3.75e-97	Length:	1647
Score:	1084.00	Matches:	209
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0

AAH26175 (1-209) x ADK67038 (1-1647)

```
Qy          1 MetLysPheLysLeuHisValAsnSerAlaArgGlnTyrLysAspLeuTrpAsnMetSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         27 ATGAAATTCAAGTTACATGTGAATTCTGCCAGGCAATACAAGGACCTGTGGAATATGAGT 86

Qy         21 AspAspLysProPheLeuCysThrAlaProGlyCysGlyGlnArgPheThrAsnGluAsp 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         87 GATGACAAACCCTTTCTATGTACTGCGCCTGGATGTGGCCAGCGTTTTACCAACGAGGAT 146

Qy         41 HisLeuAlaValHisLysHisLysHisGluMetThrLeuLysPheGlyProAlaArgAsn 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 CATTTGGCTGTCCATAAACATAAACATGAGATGACACTGAAATTTGGTCCAGCACGTAAT 206

Qy         61 AspSerValIleValAlaAspGlnThrProThrProThrArgPheLeuLysAsnCysGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 GACAGTGTCATTGTGGCTGATCAGACCCCAACACCAACAAGATTCTTGAAAAACTGTGAA 266

Qy         81 GluValGlyLeuPheAsnGluLeuAlaSerProPheGluAsnGluPheLysLysAlaSer 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 GAAGTGGGTTTGTTTAATGAGTTGGCGAGTCCATTTGAGAATGAATTCAAGAAAGCTTCA 326

Qy        101 GluAspAspIleLysLysMetProLeuAspLeuSerProLeuAlaThrProIleIleArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 GAAGATGACATTAAAAAAATGCCTCTAGATTTATCCCCTCTTGCAACACCTATCATAAGA 386

Qy        121 SerLysIleGluGluProSerValValGluThrThrHisGlnAspSerProLeuProHis 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 AGCAAAATTGAGGAGCCTTCTGTTGTAGAAACAACTCACCAGGATAGTCCTTTACCTCAC 446

Qy        141 ProGluSerThrThrSerAspGluLysGluValProLeuAlaGlnThrAlaGlnProThr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 CCAGAGTCTACTACCAGTGATGAGAAGGAAGTACCATTGGCACAAACTGCACAGCCCACA 506

Qy        161 SerAlaIleValArgProAlaSerLeuGlnValProAsnValLeuLeuThrSerSerAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 TCAGCTATTGTTCGTCCAGCATCATTACAGGTTCCCAATGTGCTGCTTACAAGTTCTGAC 566

Qy        181 SerSerValIleIleGlnGlnAlaValProSerProThrSerSerThrValIleThrGln 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 TCAAGTGTAATTATTCAGCAGGCAGTACCTTCACCAACCTCAAGTACTGTAATCACCCAG 626

Qy        201 AlaProSerSerAsnArgProIleVal 209
              ||||||||||||||||||||||||||
Db        627 GCACCATCCTCTAACAGGCCAATTGTC 653
```

RESULT 2

ADK66978
ID ADK66978 standard; DNA; 1647 BP.
XX
AC ADK66978;
XX
DT 06-MAY-2004 (first entry)
XX
DE Gene #68 for inhibitory RNA to manipulate stem cell phenotype.
XX
KW ds; gene; inhibitory RNA; RNAi; Notch signaling target gene; phenotype;
KW pluripotent stem cell.
XX
OS Homo sapiens.
XX
PN WO2003068961-A2.
XX
PD 21-AUG-2003.
XX
PF 12-FEB-2003; 2003WO-GB000579.
XX
PR 13-FEB-2002; 2002GB-00003359.
PR 13-FEB-2002; 2002GB-00003387.
XX
PA (AXOR-) AXORDIA LTD.
XX
PI Andrews P, Walsh J, Gokhale P;
XX
DR WPI; 2003-697528/66.
XX
PT New inhibitory RNA molecule having double stranded RNA molecules, useful
PT for manipulating the phenotype of stem cells, preferably pluripotential
PT stem cells.
XX
PS Disclosure; SEQ ID NO 68; 157pp; English.
XX
CC The invention relates to an inhibitory RNA (RNAi) molecule derived from a
CC nucleic acid molecule comprising a defined nucleic acid sequences given
CC in the specification or a sequence which hybridizes to the sequences and
CC encodes a Notch signaling target gene or which is a degenerate as a
CC result of the genetic code of the sequences. The methods and compositions
CC of the present invention are useful for manipulating the phenotype of
CC stem cells, preferably pluripotent stem cells. This sequence corresponds
CC to one of the nucleic acid molecules of the invention.
XX
SQ Sequence 1647 BP; 504 A; 396 C; 341 G; 406 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	3.75e-97	Length:	1647
Score:	1084.00	Matches:	209
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0

AAH26175 (1-209) x ADK66978 (1-1647)

```
Qy          1 MetLysPheLysLeuHisValAsnSerAlaArgGlnTyrLysAspLeuTrpAsnMetSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         27 ATGAAATTCAAGTTACATGTGAATTCTGCCAGGCAATACAAGGACCTGTGGAATATGAGT 86

Qy         21 AspAspLysProPheLeuCysThrAlaProGlyCysGlyGlnArgPheThrAsnGluAsp 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         87 GATGACAAACCCTTTCTATGTACTGCGCCTGGATGTGGCCAGCGTTTTACCAACGAGGAT 146

Qy         41 HisLeuAlaValHisLysHisLysHisGluMetThrLeuLysPheGlyProAlaArgAsn 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 CATTTGGCTGTCCATAAACATAAACATGAGATGACACTGAAATTTGGTCCAGCACGTAAT 206

Qy         61 AspSerValIleValAlaAspGlnThrProThrProThrArgPheLeuLysAsnCysGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 GACAGTGTCATTGTGGCTGATCAGACCCCAACACCAACAAGATTCTTGAAAAACTGTGAA 266

Qy         81 GluValGlyLeuPheAsnGluLeuAlaSerProPheGluAsnGluPheLysLysAlaSer 100
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 8, 2005, 06:13:59 ; Search time 4142 Seconds
(without alignments)
2444.988 Million cell updates/sec

Title: AAH26175
Perfect score: 1084
Sequence: 1 MKFKLHVNSARQYKDLWNMS.....PSPTSSTVITQAPSSNRPIV 209

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/AAH26175/runat_07072005_101603_18114/app_query.fasta_1.391
-DB=GenEmbl -QFMT=fastap -SUFFIX=std.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=AAH26175_@CGN_1_1_4200_@runat_07072005_101603_18114 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1084	100.0	1370	9	BC026175	BC026175 Homo sapi
2	1084	100.0	1647	6	AX821940	AX821940 Sequence
3	1084	100.0	1647	6	AX822000	AX822000 Sequence
4	1084	100.0	1647	9	HSCREBP1	X15875 Human mRNA
5	1062	98.0	1852	10	RNU38938	U38938 Rattus norv
6	1013.5	93.5	3957	10	BC042210	BC042210 Mus muscu
7	1001	92.3	1812	5	GGA17724	Y17724 Gallus gall
8	984	90.8	1464	10	AF483482	AF483482 Mus muscu
9	984	90.8	1464	10	AF483483	AF483483 Mus muscu
10	949.5	87.6	1596	10	S76657	S76657 cyclic AMP
11	940.5	86.8	3507	6	CQ851484	CQ851484 Sequence
12	940.5	86.8	3507	9	AK128731	AK128731 Homo sapi
13	937	86.4	1520	5	XLU16158	U16158 Xenopus lae
14	869	80.2	3983	9	HSM808617	BX648469 Homo sapi
15	845	78.0	1848	10	S76655	S76655 cyclic AMP
16	831.5	76.7	4130	10	BC079883	BC079883 Mus muscu
17	790	72.9	3909	10	BC082596	BC082596 Mus muscu
18	746	68.8	1209	10	RATRATF2	M65148 Rat RATF2 m
19	679.5	62.7	1302	10	S76659	S76659 cyclic AMP
20	578	53.3	2867	10	BC026483	BC026483 Mus muscu
21	569	52.5	2755	6	CQ490766	CQ490766 Sequence
22	569	52.5	2755	6	CQ496609	CQ496609 Sequence
23	569	52.5	2758	9	HSATFA	X52943 Human mRNA
24	566	52.2	3217	6	AX305250	AX305250 Sequence
25	566	52.2	3217	10	MUSTCRATF	M77167 Mouse T-cel
26	555.5	51.2	1485	9	HSATFA1	X57197 H.sapiens m
27	533	49.2	4851	12	AF050498	AF050498 Fusion tr
28	504.5	46.5	1227	6	CQ720976	CQ720976 Sequence
29	434	40.0	800	6	CQ491274	CQ491274 Sequence
30	434	40.0	800	6	CQ497144	CQ497144 Sequence
31	427	39.4	501	6	AX321198	AX321198 Sequence
32	423	39.0	787	9	BC042363	BC042363 Homo sapi
33	409.5	37.8	725	5	BX933040	BX933040 Gallus ga
34	409.5	37.8	867	5	BX930883	BX930883 Gallus ga
c 35	407.5	37.6	176033	2	AL161794	AL161794 Homo sapi
c 36	407.5	37.6	178342	9	AC007435	AC007435 Homo sapi
37	398.5	36.8	461	6	CQ727434	CQ727434 Sequence
38	390.5	36.0	2319	9	HUMCREBPAB	L05911 Human cAMP
39	390.5	36.0	2637	6	AX336935	AX336935 Sequence
40	390.5	36.0	2637	9	HUMCREBPA	L05515 Homo sapien
c 41	387.5	35.7	173793	2	AC129355	AC129355 Rattus no
c 42	386	35.6	207639	10	AL844581	AL844581 Mouse DNA
c 43	366.5	33.8	166766	2	AC119682	AC119682 Rattus no
44	332.5	30.7	432	6	CQ503391	CQ503391 Sequence
45	332.5	30.7	432	6	CQ512231	CQ512231 Sequence

ALIGNMENTS

RESULT 1
BC026175
LOCUS BC026175 1370 bp mRNA linear PRI 16-SEP-2003
DEFINITION Homo sapiens activating transcription factor 2, mRNA (cDNA clone IMAGE:4308091), complete cds.
ACCESSION BC026175
VERSION BC026175.1 GI:20072896
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1370)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,

```
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1370)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-APR-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
            cDNA Library Preparation: CLONTECH Laboratories, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Institute for Systems Biology
            http://www.systemsbiology.org
            contact: amadan@systemsbiology.org
            Anup Madan, Jessica Fahey, Erin Helton, Mark Ketteman, Anuradha
            Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 38 Row: c Column: 20
            This clone has the following problem: The cds is short compared to
            the longest cds in the locus.
FEATURES             Location/Qualifiers
     source          1. .1370
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4308091"
                     /tissue_type="Skin, melanotic melanoma, high MDR."
                     /clone_lib="NIH_MGC_62"
                     /lab_host="DH10B"
                     /note="Vector: pDNR-LIB"
     gene            1. .1370
                     /gene="ATF2"
                     /note="synonyms: CRE-BP1, TREB7, HB16"
                     /db_xref="LocusID:1386"
                     /db_xref="MIM:123811"
     CDS             267. .896
                     /gene="ATF2"
                     /codon_start=1
                     /product="ATF2 protein"
                     /protein_id="AAH26175.1"
                     /db_xref="GI:20072897"
                     /db_xref="LocusID:1386"
                     /db_xref="MIM:123811"
                     /translation="MKFKLHVNSARQYKDLWNMSDDKPFLCTAPGCGQRFTNEDHLAV
                     HKHKHEMTLKFGPARNDSVIVADQTPTPTRFLKNCEEVGLFNELASPFENEFKKASED
                     DIKKMPLDLSPLATPIIRSKIEEPSVVETTHQDSPLPHPESTTSDEKEVPLAQTAQPT
                     SAIVRPASLQVPNVLLTSSDSSVIIQQAVPSPTSSTVITQAPSSNRPIV"
ORIGIN

Alignment Scores:
Pred. No.:              8.56e-76        Length:         1370
Score:                  1084.00         Matches:        209
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     9               Gaps:           0

AAH26175 (1-209) x BC026175 (1-1370)

Qy          1 MetLysPheLysLeuHisValAsnSerAlaArgGlnTyrLysAspLeuTrpAsnMetSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 ATGAAATTCAAGTTACATGTGAATTCTGCCAGGCAATACAAGGACCTGTGGAATATGAGT 326

Qy         21 AspAspLysProPheLeuCysThrAlaProGlyCysGlyGlnArgPheThrAsnGluAsp 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 GATGACAAACCCTTTCTATGTACTGCGCCTGGATGTGGCCAGCGTTTTACCAACGAGGAT 386

Qy         41 HisLeuAlaValHisLysHisLysHisGluMetThrLeuLysPheGlyProAlaArgAsn 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 CATTTGGCTGTCCATAAACATAAACATGAGATGACACTGAAATTTGGTCCAGCACGTAAT 446

Qy         61 AspSerValIleValAlaAspGlnThrProThrProThrArgPheLeuLysAsnCysGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 GACAGTGTCATTGTGGCTGATCAGACCCCAACACCAACAAGATTCTTGAAAAACTGTGAA 506

Qy         81 GluValGlyLeuPheAsnGluLeuAlaSerProPheGluAsnGluPheLysLysAlaSer 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 GAAGTGGGTTTGTTTAATGAGTTGGCGAGTCCATTTGAGAATGAATTCAAGAAAGCTTCA 566

Qy        101 GluAspAspIleLysLysMetProLeuAspLeuSerProLeuAlaThrProIleIleArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 GAAGATGACATTAAAAAAATGCCTCTAGATTTATCCCCTCTTGCAACACCTATCATAAGA 626

Qy        121 SerLysIleGluGluProSerValValGluThrThrHisGlnAspSerProLeuProHis 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 AGCAAAATTGAGGAGCCTTCTGTTGTAGAAACAACTCACCAGGATAGTCCTTTACCTCAC 686

Qy        141 ProGluSerThrThrSerAspGluLysGluValProLeuAlaGlnThrAlaGlnProThr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 CCAGAGTCTACTACCAGTGATGAGAAGGAAGTACCATTGGCACAAACTGCACAGCCCACA 746

Qy        161 SerAlaIleValArgProAlaSerLeuGlnValProAsnValLeuLeuThrSerSerAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 TCAGCTATTGTTCGTCCAGCATCATTACAGGTTCCCAATGTGCTGCTTACAAGTTCTGAC 806

Qy        181 SerSerValIleIleGlnGlnAlaValProSerProThrSerSerThrValIleThrGln 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        807 TCAAGTGTAATTATTCAGCAGGCAGTACCTTCACCAACCTCAAGTACTGTAATCACCCAG 866

Qy        201 AlaProSerSerAsnArgProIleVal 209
              ||||||||||||||||||||||||||
Db        867 GCACCATCCTCTAACAGGCCAATTGTG 893

RESULT 2
AX821940
LOCUS       AX821940               1647 bp    DNA     linear   PAT 10-DEC-2003
DEFINITION  Sequence 68 from Patent WO03068961.
ACCESSION   AX821940
VERSION     AX821940.1  GI:39725161
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Andrews,P.A., Walsh,J.A. and Gokhale,P.A.
  TITLE     Method to modify differentiation of pluripotential stem cells
  JOURNAL   Patent: WO 03068961-A 68 21-AUG-2003;
            Axordia Limited (GB)
FEATURES             Location/Qualifiers
     source          1. .1647
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              1.05e-75        Length:         1647
Score:                  1084.00         Matches:        209
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 8, 2005, 03:17:01 ; Search time 90 Seconds
(without alignments)
897.145 Million cell updates/sec

Title: AAH26175
Perfect score: 1084
Sequence: 1 MKFKLHVNSARQYKDLWNMS.....PSPTSSTVITQAPSSNRPIV 209

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1726216 seqs, 386330316 residues

Total number of hits satisfying chosen parameters: 1726216

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	156	14.4	29	9	US-09-970-515-20	Sequence 20, Appl
2	156	14.4	29	14	US-10-165-250A-20	Sequence 20, Appl
3	156	14.4	29	15	US-10-342-683-20	Sequence 20, Appl
4	156	14.4	29	15	US-10-457-614A-20	Sequence 20, Appl
5	156	14.4	29	17	US-10-924-028-20	Sequence 20, Appl
6	144	13.3	27	13	US-10-115-178-6	Sequence 6, Appli
7	115.5	10.7	830	16	US-10-425-115-298770	Sequence 298770,
8	112	10.3	447	16	US-10-723-860-177	Sequence 177, App
9	111.5	10.3	991	16	US-10-425-115-231754	Sequence 231754,
10	110.5	10.2	1292	17	US-10-741-600-1572	Sequence 1572, Ap
11	110.5	10.2	1367	9	US-09-801-368-108	Sequence 108, App
12	110	10.1	730	16	US-10-425-115-299626	Sequence 299626,
13	108	10.0	992	16	US-10-425-115-297854	Sequence 297854,
14	106	9.8	443	14	US-10-244-367-42	Sequence 42, Appl
15	105.5	9.7	885	16	US-10-425-115-231789	Sequence 231789,
16	105	9.7	960	16	US-10-425-115-298810	Sequence 298810,
17	105	9.7	3394	16	US-10-408-765A-529	Sequence 529, App
18	104.5	9.6	700	14	US-10-171-311-160	Sequence 160, App
19	104.5	9.6	700	14	US-10-301-822-141	Sequence 141, App
20	104.5	9.6	700	16	US-10-751-736-103	Sequence 103, App
21	104.5	9.6	700	17	US-10-645-756-36	Sequence 36, Appl
22	104.5	9.6	777	16	US-10-425-115-298589	Sequence 298589,
23	104	9.6	894	16	US-10-425-115-231793	Sequence 231793,
24	104	9.6	1983	16	US-10-723-860-1890	Sequence 1890, Ap
25	103	9.5	492	16	US-10-425-115-298835	Sequence 298835,
26	103	9.5	1166	15	US-10-104-047-2949	Sequence 2949, Ap
27	102.5	9.5	533	15	US-10-373-238-17	Sequence 17, Appl
28	102	9.4	1689	16	US-10-425-115-300468	Sequence 300468,
29	102	9.4	1845	17	US-10-732-923-8645	Sequence 8645, Ap
30	101.5	9.4	619	16	US-10-739-930-9129	Sequence 9129, Ap
31	101.5	9.4	999	16	US-10-425-115-231720	Sequence 231720,
32	101	9.3	532	9	US-09-833-790-428	Sequence 428, App
33	101	9.3	533	16	US-10-761-169-22	Sequence 22, Appl
34	101	9.3	878	16	US-10-722-050-2	Sequence 2, Appli
35	101	9.3	907	9	US-09-328-599A-1	Sequence 1, Appli
36	101	9.3	910	16	US-10-425-115-299670	Sequence 299670,
37	100.5	9.3	501	14	US-10-244-367-44	Sequence 44, Appl
38	100.5	9.3	1087	16	US-10-425-115-300675	Sequence 300675,
39	100	9.2	97	9	US-09-925-796-15	Sequence 15, Appl
40	100	9.2	97	9	US-09-941-450-15	Sequence 15, Appl
41	100	9.2	97	9	US-09-942-090-15	Sequence 15, Appl
42	100	9.2	97	16	US-10-843-944-15	Sequence 15, Appl
43	100	9.2	97	17	US-10-922-546-15	Sequence 15, Appl
44	100	9.2	267	9	US-09-864-761-37826	Sequence 37826, A
45	100	9.2	1557	10	US-09-940-316B-27	Sequence 27, Appl

ALIGNMENTS

RESULT 1
US-09-970-515-20
; Sequence 20, Application US/09970515
; Patent No. US20020127676A1
; GENERAL INFORMATION:
; APPLICANT: Bonny, Christophe
; TITLE OF INVENTION: CELL-PERMEABLE PEPTIDE INHIBITORS OF THE JNK SIGNAL
; TITLE OF INVENTION: TRANSDUCTION PATHWAY
; FILE REFERENCE: 20349-501 DIV
; CURRENT APPLICATION NUMBER: US/09/970,515
; CURRENT FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: 09/503,954
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: USSN 60/158,774
; PRIOR FILING DATE: 1999-10-12
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically synthesized
US-09-970-515-20

Query Match 14.4%; Score 156; DB 9; Length 29;
Best Local Similarity 100.0%; Pred. No. 3.2e-07;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 37 TNEDHLAVHKHKHEMTLKFGPARNDSVIV 65
||||||||||||||||||||||||||||
Db 1 TNEDHLAVHKHKHEMTLKFGPARNDSVIV 29

```
RESULT 2
US-10-165-250A-20
; Sequence 20, Application US/10165250A
; Publication No. US20030108539A1
; GENERAL INFORMATION:
;  APPLICANT: Bonny, Christophe
;  TITLE OF INVENTION: CELL-PERMEABLE PEPTIDE INHIBITORS OF THE JNK SIGNAL
;  TITLE OF INVENTION:  TRANSDUCTION PATHWAY
;  FILE REFERENCE: 20349-501 CIP
;  CURRENT APPLICATION NUMBER: US/10/165,250A
;  CURRENT FILING DATE:  2002-06-07
;  PRIOR APPLICATION NUMBER: 09/503,954
;  PRIOR FILING DATE: 2000-02-14
;  PRIOR APPLICATION NUMBER: USSN 60/158,774
;  PRIOR FILING DATE: 1999-10-12
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
;   LENGTH: 29
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically synthesized
US-10-165-250A-20

  Query Match             14.4%;  Score 156;  DB 14;  Length 29;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-07;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         37 TNEDHLAVHKHKHEMTLKFGPARNDSVIV 65
              |||||||||||||||||||||||||||||
Db          1 TNEDHLAVHKHKHEMTLKFGPARNDSVIV 29


RESULT 3
US-10-342-683-20
; Sequence 20, Application US/10342683
; Publication No. US20030220480A1
; GENERAL INFORMATION:
;  APPLICANT: Bonny, Christophe
;  TITLE OF INVENTION: CELL-PERMEABLE PEPTIDE INHIBITORS OF THE JNK SIGNAL
;  TITLE OF INVENTION:  TRANSDUCTION PATHWAY
;  FILE REFERENCE: 20349-501
;  CURRENT APPLICATION NUMBER: US/10/342,683
;  CURRENT FILING DATE:  2003-01-15
;  PRIOR APPLICATION NUMBER: US/09/503,954A
;  PRIOR FILING DATE: 2000-02-14
;  PRIOR APPLICATION NUMBER: USSN 60/158,774
;  PRIOR FILING DATE: 1999-10-12
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
;   LENGTH: 29
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically synthesized
US-10-342-683-20

  Query Match             14.4%;  Score 156;  DB 15;  Length 29;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-07;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         37 TNEDHLAVHKHKHEMTLKFGPARNDSVIV 65
              |||||||||||||||||||||||||||||
Db          1 TNEDHLAVHKHKHEMTLKFGPARNDSVIV 29


RESULT 4
US-10-457-614A-20
; Sequence 20, Application US/10457614A
; Publication No. US20040082509A1
; GENERAL INFORMATION:
;  APPLICANT: Bonny, Christope
;  TITLE OF INVENTION: CELL-PERMEABLE PEPTIDE INHIBITORS OF THE JNK SIGNAL
;  TITLE OF INVENTION:  TRANSDUCTION PATHWAY
;  FILE REFERENCE: 20349-501CIP2
;  CURRENT APPLICATION NUMBER: US/10/457,614A
;  CURRENT FILING DATE:  2003-06-09
;  PRIOR APPLICATION NUMBER: 60/158,774
;  PRIOR FILING DATE: 1999-10-12
;  PRIOR APPLICATION NUMBER: 09/503,954
;  PRIOR FILING DATE: 2000-02-14
;  PRIOR APPLICATION NUMBER: 10/165,250
;  PRIOR FILING DATE: 2002-06-07
;  PRIOR APPLICATION NUMBER: 60/347,062
;  PRIOR FILING DATE: 2002-01-09
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
;   LENGTH: 29
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically synthesized
US-10-457-614A-20

  Query Match             14.4%;  Score 156;  DB 15;  Length 29;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-07;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         37 TNEDHLAVHKHKHEMTLKFGPARNDSVIV 65
              |||||||||||||||||||||||||||||
Db          1 TNEDHLAVHKHKHEMTLKFGPARNDSVIV 29


RESULT 5
US-10-924-028-20
; Sequence 20, Application US/10924028
; Publication No. US20050043241A1
; GENERAL INFORMATION:
;  APPLICANT: Bonny, Christophe
;  TITLE OF INVENTION: CELL-PERMEABLE PEPTIDE INHIBITORS OF THE JNK SIGNAL
;  TITLE OF INVENTION:  TRANSDUCTION PATHWAY
;  FILE REFERENCE: 20349-501 DIV
;  CURRENT APPLICATION NUMBER: US/10/924,028
;  CURRENT FILING DATE:  2004-08-23
;  PRIOR APPLICATION NUMBER: US/09/970,515
;  PRIOR FILING DATE: 2001-10-03
;  PRIOR APPLICATION NUMBER: 09/503,954
;  PRIOR FILING DATE: 2000-02-14
;  PRIOR APPLICATION NUMBER: USSN 60/158,774
;  PRIOR FILING DATE: 1999-10-12
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
;   LENGTH: 29
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically synthesized
US-10-924-028-20

  Query Match             14.4%;  Score 156;  DB 17;  Length 29;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-07;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         37 TNEDHLAVHKHKHEMTLKFGPARNDSVIV 65
              |||||||||||||||||||||||||||||
Db          1 TNEDHLAVHKHKHEMTLKFGPARNDSVIV 29


RESULT 6
US-10-115-178-6
; Sequence 6, Application US/10115178
; Publication No. US20020119135A1
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 8, 2005, 03:11:51 ; Search time 33 Seconds
(without alignments)
472.777 Million cell updates/sec

Title: AAH26175
Perfect score: 1084
Sequence: 1 MKFKLHVNSARQYKDLWNMS.....PSPTSSTVITQAPSSNRPIV 209

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1084	100.0	501	4	US-09-949-016-11281	Sequence 11281, A
2	257	23.7	52	4	US-09-513-999C-7023	Sequence 7023, Ap
3	156	14.4	29	4	US-09-503-954A-20	Sequence 20, Appl
4	156	14.4	29	4	US-09-970-515A-20	Sequence 20, Appl
5	144	13.3	27	3	US-08-819-177-6	Sequence 6, Appli
6	112	10.3	447	3	US-09-234-332-11	Sequence 11, Appl
7	106	9.8	443	4	US-09-342-325C-42	Sequence 42, Appl
8	104.5	9.6	700	4	US-09-976-594-67	Sequence 67, Appl
9	104.5	9.6	700	4	US-09-538-092-893	Sequence 893, App
10	102.5	9.5	496	2	US-08-224-482-2	Sequence 2, Appli
11	102.5	9.5	533	1	US-08-040-548-1	Sequence 1, Appli
12	102.5	9.5	533	1	US-08-466-344-1	Sequence 1, Appli
13	102.5	9.5	533	4	US-09-546-013-21	Sequence 21, Appl
14	102.5	9.5	533	6	5206152-2	Patent No. 5206152
15	102.5	9.5	533	6	5206152-2	Patent No. 5206152
16	101	9.3	878	4	US-09-556-706B-2	Sequence 2, Appli
17	101	9.3	878	4	US-09-724-418A-2	Sequence 2, Appli
18	101	9.3	907	3	US-08-783-774-2	Sequence 2, Appli
19	101	9.3	907	4	US-09-328-599A-1	Sequence 1, Appli
20	101	9.3	907	5	PCT-US95-04611A-19	Sequence 19, Appl
21	100.5	9.3	501	4	US-09-342-325C-44	Sequence 44, Appl
22	100	9.2	97	4	US-09-395-448-15	Sequence 15, Appl
23	100	9.2	97	4	US-09-925-796-15	Sequence 15, Appl
24	100	9.2	97	4	US-09-941-450-15	Sequence 15, Appl
25	100	9.2	1557	4	US-09-410-551B-27	Sequence 27, Appl
26	100	9.2	1557	4	US-09-940-316B-27	Sequence 27, Appl
27	97.5	9.0	306	4	US-10-029-180-60	Sequence 60, Appl
28	96	8.9	522	4	US-09-949-016-11145	Sequence 11145, A
29	96	8.9	667	4	US-09-949-016-7759	Sequence 7759, Ap
30	96	8.9	753	4	US-09-949-016-6676	Sequence 6676, Ap
31	95.5	8.8	556	4	US-09-949-016-7134	Sequence 7134, Ap
32	95.5	8.8	613	4	US-09-438-833-6	Sequence 6, Appli
33	95	8.8	1140	4	US-09-538-092-647	Sequence 647, App
34	94.5	8.7	311	4	US-09-438-833-7	Sequence 7, Appli
35	94.5	8.7	532	4	US-09-949-016-7389	Sequence 7389, Ap
36	94.5	8.7	652	4	US-09-438-833-5	Sequence 5, Appli
37	94.5	8.7	756	4	US-09-438-833-11	Sequence 11, Appl
38	94.5	8.7	805	2	US-08-480-473B-4	Sequence 4, Appli
39	94.5	8.7	805	3	US-08-915-213-4	Sequence 4, Appli
40	94.5	8.7	805	3	US-09-235-217-4	Sequence 4, Appli
41	94.5	8.7	805	5	PCT-US96-10251-4	Sequence 4, Appli
42	94.5	8.7	813	4	US-09-438-833-12	Sequence 12, Appl
43	94.5	8.7	826	1	US-08-785-241-6	Sequence 6, Appli
44	94.5	8.7	826	2	US-08-480-473B-2	Sequence 2, Appli
45	94.5	8.7	826	3	US-08-915-213-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-949-016-11281
; Sequence 11281, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11281
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11281

Query Match 100.0%; Score 1084; DB 4; Length 501;
Best Local Similarity 100.0%; Pred. No. 1.1e-106;
Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MKFKLHVNSARQYKDLWNMSDDKPFLCTAPGCGQRFTNEDHLAVHKHKHEMTLKFGPARN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5 MKFKLHVNSARQYKDLWNMSDDKPFLCTAPGCGQRFTNEDHLAVHKHKHEMTLKFGPARN 64

Qy         61 DSVIVADQTPTPTRFLKNCEEVGLFNELASPFENEFKKASEDDIKKMPLDLSPLATPIIR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 DSVIVADQTPTPTRFLKNCEEVGLFNELASPFENEFKKASEDDIKKMPLDLSPLATPIIR 124

Qy        121 SKIEEPSVVETTHQDSPLPHPESTTSDEKEVPLAQTAQPTSAIVRPASLQVPNVLLTSSD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 SKIEEPSVVETTHQDSPLPHPESTTSDEKEVPLAQTAQPTSAIVRPASLQVPNVLLTSSD 184

Qy        181 SSVIIQQAVPSPTSSTVITQAPSSNRPIV 209
              |||||||||||||||||||||||||||||
Db        185 SSVIIQQAVPSPTSSTVITQAPSSNRPIV 213
```

RESULT 2
US-09-513-999C-7023
; Sequence 7023, Application US/09513999C
; Patent No. 6783961

```
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 7023
;   LENGTH: 52
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: 6
;   OTHER INFORMATION: Xaa=Pro or Ser
US-09-513-999C-7023

  Query Match             23.7%;  Score 257;  DB 4;  Length 52;
  Best Local Similarity   86.5%;  Pred. No. 3.9e-20;
  Matches   45;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         19 MSDDKPFLCTAPGCGQRFTNEDHLAVHKHKHEMTLKFGPARNDSVIVADQTP 70
              | ||: |:| |||||||||||||||||||||||||||||| ||||:|||||
Db          1 MGDDRXFVCNAPGCGQRFTNEDHLAVHKHKHEMTLKFGPARTDSVIIADQTP 52


RESULT 3
US-09-503-954A-20
; Sequence 20, Application US/09503954A
; Patent No. 6610820
; GENERAL INFORMATION:
;  APPLICANT: Bonny, Christophe
;  TITLE OF INVENTION: CELL-PERMEABLE PEPTIDE INHIBITORS OF THE JNK SIGNAL
;  TITLE OF INVENTION:  TRANSDUCTION PATHWAY
;  FILE REFERENCE: 20349-501
;  CURRENT APPLICATION NUMBER: US/09/503,954A
;  CURRENT FILING DATE:  2000-02-14
;  PRIOR APPLICATION NUMBER: USSN 60/158,774
;  PRIOR FILING DATE: 1999-10-12
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
;   LENGTH: 29
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically synthesized
US-09-503-954A-20

  Query Match             14.4%;  Score 156;  DB 4;  Length 29;
  Best Local Similarity   100.0%;  Pred. No. 9.1e-10;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         37 TNEDHLAVHKHKHEMTLKFGPARNDSVIV 65
              |||||||||||||||||||||||||||||
Db          1 TNEDHLAVHKHKHEMTLKFGPARNDSVIV 29


RESULT 4
US-09-970-515A-20
; Sequence 20, Application US/09970515A
; Patent No. 6780970
; GENERAL INFORMATION:
;  APPLICANT: Bonny, Christophe
;  TITLE OF INVENTION: CELL-PERMEABLE PEPTIDE INHIBITORS OF THE JNK SIGNAL
;  TITLE OF INVENTION:  TRANSDUCTION PATHWAY
;  FILE REFERENCE: 20349-501 DIV
;  CURRENT APPLICATION NUMBER: US/09/970,515A
;  CURRENT FILING DATE:  2001-10-03
;  PRIOR APPLICATION NUMBER: 09/503,954
;  PRIOR FILING DATE: 2000-02-14
;  PRIOR APPLICATION NUMBER: 60/158,774
;  PRIOR FILING DATE: 1999-10-12
;  NUMBER OF SEQ ID NOS: 21
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
;   LENGTH: 29
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically synthesized
US-09-970-515A-20

  Query Match             14.4%;  Score 156;  DB 4;  Length 29;
  Best Local Similarity   100.0%;  Pred. No. 9.1e-10;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         37 TNEDHLAVHKHKHEMTLKFGPARNDSVIV 65
              |||||||||||||||||||||||||||||
Db          1 TNEDHLAVHKHKHEMTLKFGPARNDSVIV 29


RESULT 5
US-08-819-177-6
; Sequence 6, Application US/08819177
; Patent No. 6043083
;  GENERAL INFORMATION:
;    APPLICANT:  Davis, Roger J.
;    APPLICANT:  Dickens, Martin
;    TITLE OF INVENTION:  INHIBITORS OF THE JNK SIGNAL
;    TITLE OF INVENTION:  TRANSDUCTION PATHWAY AND METHODS OF USE
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson, P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows95
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/819,177
;      FILING DATE:  28 April 1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fasse, J. Peter
;      REGISTRATION NUMBER:  32,983
;      REFERENCE/DOCKET NUMBER:  07917/037001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-542-5070
;      TELEFAX:  617-542-8906
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  27 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      OTHER INFORMATION:  ATF-2 JNK-binding domain
US-08-819-177-6

  Query Match             13.3%;  Score 144;  DB 3;  Length 27;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-08;
  Matches   27;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 8, 2005, 03:06:25 ; Search time 96 Seconds
(without alignments)
842.010 Million cell updates/sec

Title: AAH26175
Perfect score: 1084
Sequence: 1 MKFKLHVNSARQYKDLWNMS.....PSPTSSTVITQAPSSNRPIV 209

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	940.5	86.8	482	8	ADR10403	Adr10403 Human pro
2	571	52.7	413	8	ADP80860	Adp80860 Mouse C13
3	566	52.2	411	5	ABB57020	Abb57020 Mouse isc
4	390.5	36.0	501	7	ADC18733	Adc18733 Human cyc
5	257	23.7	52	3	AAG02942	Aag02942 Human sec
6	171	15.8	359	4	ABB67482	Abb67482 Drosophil
7	156	14.4	29	4	AAB68325	Aab68325 c-Jun ami
8	156	14.4	29	7	ADF17925	Adf17925 Synthetic
9	156	14.4	29	8	ADO01349	Ado01349 c-Jun ami
10	112	10.3	447	3	AAB14349	Aab14349 Human Zic
11	112	10.3	447	8	ADN04236	Adn04236 Antipsori
12	112	10.3	447	8	ADQ17360	Adq17360 Human sof
13	110.5	10.2	1292	8	ADQ39909	Adq39909 Human myo
14	110	10.1	447	5	ABB57042	Abb57042 Mouse isc
15	109	10.1	752	4	ABB60124	Abb60124 Drosophil
16	107.5	9.9	791	4	ABB67268	Abb67268 Drosophil
17	107.5	9.9	845	4	ABB64280	Abb64280 Drosophil
18	106.5	9.8	1292	7	ADC31295	Adc31295 Human nov
19	106.5	9.8	1484	4	ABG18797	Abg18797 Novel hum
20	105	9.7	2992	8	ADP30190	Adp30190 Human sec
21	105	9.7	3065	8	ADP30259	Adp30259 Human sec
22	105	9.7	3394	7	ADJ68723	Adj68723 Human hea
23	104.5	9.6	512	8	ADP22648	Adp22648 Sea-squir
24	104.5	9.6	700	4	AAM40233	Aam40233 Human pol
25	104.5	9.6	700	6	ABR92125	Abr92125 Human cer
26	104.5	9.6	700	7	ADD14182	Add14182 Human src
27	104.5	9.6	700	7	ADN95317	Adn95317 Human BEC
28	104.5	9.6	700	8	ADL12338	Adl12338 Human ste
29	104.5	9.6	700	8	ADL70593	Adl70593 Cervical
30	104.5	9.6	700	8	ABM82242	Abm82242 Tumour-as
31	104.5	9.6	724	4	AAM42019	Aam42019 Human pol
32	104	9.6	1982	4	ABG16404	Abg16404 Novel hum
33	104	9.6	1982	4	ABG19656	Abg19656 Novel hum
34	104	9.6	1983	8	ADQ19071	Adq19071 Human sof
35	104	9.6	2768	4	ABB68397	Abb68397 Drosophil
36	103.5	9.5	929	8	ADR08894	Adr08894 Human pro
37	103.5	9.5	1483	8	ADQ97206	Adq97206 Mouse can
38	103	9.5	1044	5	ABP64908	Abp64908 Human pro
39	103	9.5	1166	7	ADB64795	Adb64795 Human pro
40	103	9.5	4873	6	ABO14747	Abo14747 Novel hum
41	102.5	9.5	442	7	ADE63885	Ade63885 Rat Prote
42	102.5	9.5	496	2	AAW76984	Aaw76984 Mouse Egr
43	102.5	9.5	533	1	AAP93113	Aap93113 Egr-1. 9/
44	102.5	9.5	533	2	AAR63129	Aar63129 Mouse Egr
45	102.5	9.5	533	4	AAG78054	Aag78054 Human zin

ALIGNMENTS

RESULT 1
ADR10403
ID ADR10403 standard; protein; 482 AA.
XX
AC ADR10403;
XX
DT 04-NOV-2004 (first entry)
XX
DE Human protein useful for treating neurological disease Seq 3909.
XX
KW human; oligo-capping method; diagnostic marker; gene therapy;
KW osteoporosis; neurological disease; Alzheimer's disease;
KW Parkinson's disease; dementia; short memory; cancer;
KW sense or motor function; emotional reaction; fear response; panic;
KW osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;
KW tranquiliser.
XX
OS Homo sapiens.
XX
PN EP1447413-A2.
XX
PD 18-AUG-2004.
XX
PF 12-FEB-2004; 2004EP-00003145.
XX
PR 14-FEB-2003; 2003JP-00102207.
PR 09-MAY-2003; 2003JP-00131452.
XX
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Isogai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T;
PI Wakamatsu A, Ishii S, Nagai K, Irie R;
XX
DR WPI; 2004-583265/57.
DR N-PSDB; ADR08447.
XX
PT New 1995 cDNA, useful for treating osteoporosis, neurological diseases,
PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX
PS Claim 1; SEQ ID NO 3909; 2686pp; English.
XX
CC This invention relates to novel, isolated full length human cDNA
CC molecules and the encoded proteins thereof. Specifically, it refers to
CC cDNA clones obtained by an oligo-capping method, where none of these
CC clones are identical to any known human mRNAs. The present invention
CC describes an immunoassay to identify agonists and antagonists, as well as
CC antibodies, antisense molecules and siRNAs that can all be used to bind
CC to and modulate expression of the cDNA molecules. As such, these

```
CC   molecules are useful for diagnostic markers or therapeutic targets for
CC   the various diseases or morbid states. In particular, they are useful in
CC   gene therapy for treating osteoporosis, neurological disease, Alzheimer's
CC   disease, Parkinson's disease, dementia, short memory and various cancers,
CC   as well as for maintaining equilibrium of sense or motor function, and
CC   for treating emotional reaction, fear response and panic. Accordingly,
CC   they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian,
CC   cytostatic and tranquiliser activities. This polypeptide is a protein
CC   encoded by a full length human cDNA sequence of the invention. NOTE: This
CC   sequence is not given in the sequence listing of the specification but
CC   can be obtained on CD-ROM from the European Patent Office, Vienna Sub-
CC   office.
XX
SQ   Sequence 482 AA;

  Query Match             86.8%;  Score 940.5;  DB 8;  Length 482;
  Best Local Similarity   89.0%;  Pred. No. 2.8e-79;
  Matches  186;  Conservative    0;  Mismatches    0;  Indels   23;  Gaps    1;

Qy          1 MKFKLHVNSARQYKDLWNMSDDKPFLCTAPGCGQRFTNEDHLAVHKHKHEMTLKFGPARN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKFKLHVNSARQYKDLWNMSDDKPFLCTAPGCGQRFTNEDHL------------------ 42

Qy         61 DSVIVADQTPTPTRFLKNCEEVGLFNELASPFENEFKKASEDDIKKMPLDLSPLATPIIR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         43 -----ADQTPTPTRFLKNCEEVGLFNELASPFENEFKKASEDDIKKMPLDLSPLATPIIR 97

Qy        121 SKIEEPSVVETTHQDSPLPHPESTTSDEKEVPLAQTAQPTSAIVRPASLQVPNVLLTSSD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         98 SKIEEPSVVETTHQDSPLPHPESTTSDEKEVPLAQTAQPTSAIVRPASLQVPNVLLTSSD 157

Qy        181 SSVIIQQAVPSPTSSTVITQAPSSNRPIV 209
              |||||||||||||||||||||||||||||
Db        158 SSVIIQQAVPSPTSSTVITQAPSSNRPIV 186


RESULT 2
ADP80860
ID   ADP80860 standard; protein; 413 AA.
XX
AC   ADP80860;
XX
DT   09-SEP-2004  (first entry)
XX
DE   Mouse C130020M04Rik amino acid sequence SEQ ID NO:85.
XX
KW   c-fos; c-fos interacting protein; fos interacting protein chromosome X;
KW   Fip-cx; screening; Fc-fos interacting protein inhibitor; mouse.
XX
OS   Mus musculus.
XX
PN   WO2004053121-A1.
XX
PD   24-JUN-2004.
XX
PF   19-NOV-2003; 2003WO-JP014749.
XX
PR   11-DEC-2002; 2002JP-00360046.
XX
PA   (UYKE-) UNIV KEIO.
XX
PI   Miyamoto E,  Ishizaka M,  Yanagawa H;
XX
DR   WPI; 2004-517250/49.
DR   N-PSDB; ADP80924.
XX
PT   New proteins that interact with fos, e.g., fos interacting protein
PT   chromosome X (Fip-cx).
XX
PS   Claim 57; SEQ ID NO 85; 192pp; Japanese.
XX
CC   The present invention describes a protein (I) that interacts with c-fos
CC   (e.g., fos interacting protein chromosome X (Fip-cx), Fip-cx.1, fos
CC   interacting protein chromosome ex.2, or fos interacting protein
CC   chromosome 4). Also described: (1) a nucleic acid (II) that encodes (I);
CC   (2) an inhibitor (III) that inhibits the interaction of (I) encoded by
CC   (II) with c-fos protein; (3) an inhibitor (IV) that inhibits interaction
CC   of a protein with c-fos protein; and (4) detecting (M1) the interaction
CC   of a protein as mentioned in (3) with c-fos protein, by contacting the
CC   protein with c-fos, to form a composite. (I) is useful for detecting its
CC   interaction with c-fos, to form a complex. (M1) is useful for screening
CC   the protein that interacts with c-fos which involves performing the
CC   detection process and selecting the protein that interacts with c-fos.
CC   (M1) is useful for screening a protein that interacts with c-fos which
CC   involves performing (M1) and selecting the detected protein. (I) is
CC   useful for screening inhibitors that interact with c-fos. The present
CC   sequence represents a mouse C130020M04Rik amino acid sequence, which can
CC   interact with c-fos in the exemplification of the present invention.
XX
SQ   Sequence 413 AA;

  Query Match             52.7%;  Score 571;  DB 8;  Length 413;
  Best Local Similarity   63.8%;  Pred. No. 9.6e-45;
  Matches  125;  Conservative   16;  Mismatches   39;  Indels   16;  Gaps    6;

Qy         19 MSDDKPFLCTAPGCGQRFTNEDHLAVHKHKHEMTLKFGPARNDSVIVADQTPTPTRFLKN 78
              | ||:||:|:|||||||||||||||||||||||||||||||| ||||:||||||||||||
Db          1 MGDDRPFVCSAPGCGQRFTNEDHLAVHKHKHEMTLKFGPARTDSVIIADQTPTPTRFLKN 60

Qy         79 CEEVGLFNELASPFENEFKKASEDDIKK---MPLDLSPLATPIIRSKIEEPSVVETTHQD 135
              |||||||||||| ||:||||||:|| ||    |||:|  :|| |: | |||  |:::  |
Db         61 CEEVGLFNELASSFEHEFKKASDDDEKKGAAGPLDMSLPSTPDIKIKEEEPVEVDSSPPD 120

Qy        136 SPLPHPESTTSDEKEV---PLAQTAQPTSAIVRPASLQVPNVLLTSSDSSVIIQQAVPSP 192
              ||   | |    ||||   |:   : ||  |||| ||      |     |    :   :|||
Db        121 SPASSPCSPPLKEKEVTTKPVV-ISTPTPTIVRPGSLP----LHLGYDP---LHPTLPSP 172

Qy        193 TSSTVITQAPSSNRPI 208
              ||  |||||| ||| |
Db        173 TS--VITQAPPSNRQI 186


RESULT 3
ABB57020
ID   ABB57020 standard; protein; 411 AA.
XX
AC   ABB57020;
XX
DT   07-MAR-2002  (first entry)
XX
DE   Mouse ischaemic condition related protein sequence SEQ ID NO:2.
XX
KW   Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
KW   vasospastic ischaemia; ischaemic condition; ischaemic disease.
XX
OS   Mus musculus.
XX
PN   WO200188188-A2.
XX
PD   22-NOV-2001.
XX
PF   18-MAY-2001; 2001WO-JP004192.
XX
PR   18-MAY-2000; 2000JP-00145977.
XX
PA   (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX
PI   Ishikawa K,  Asai S,  Takahashi Y,  Nagata T,  Ishii Y;
XX
DR   WPI; 2002-034733/04.
DR   N-PSDB; ABI99202.
XX
PT   Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT   expression levels of particular genes defined in the specification or by
PT   determining the expression profile of a gene group comprising these
PT   genes.
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 8, 2005, 11:19:57 ; Search time 91 Seconds
(without alignments)
1176.094 Million cell updates/sec

Title: AAH26175
Perfect score: 209
Sequence: 1 MKFKLHVNSARQYKDLWNMS.....PSPTSSTVITQAPSSNRPIV 209

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1612378 seqs, 512079187 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	209	100.0	209	2	Q8TAR1	Q8tar1 homo sapien
2	191	91.4	234	2	Q8BN75	Q8bn75 mus musculu
3	191	91.4	420	2	Q8CBR9	Q8cbr9 mus musculu
4	191	91.4	487	1	ATF2_HUMAN	P15336 homo sapien
5	191	91.4	487	1	ATF2_MOUSE	P16951 mus musculu
6	191	91.4	487	2	Q8CGB4	Q8cgb4 mus musculu
7	143	68.4	440	2	Q68FE3	Q68fe3 mus musculu
8	127	60.8	487	1	ATF2_RAT	Q00969 rattus norv
9	102	48.8	487	1	ATF2_CHICK	O93602 gallus gall
10	62	29.7	486	2	Q91576	Q91576 xenopus lae
11	34	16.3	144	2	Q96JT8	Q96jt8 homo sapien
12	31	14.8	307	2	Q9UD83	Q9ud83 homo sapien
13	31	14.8	413	2	Q8R0S1	Q8r0s1 mus musculu
14	31	14.8	494	1	ATF7_HUMAN	P17544 homo sapien
15	29	13.9	117	2	Q8IVR8	Q8ivr8 homo sapien
16	19	9.1	135	2	Q8BKN9	Q8bkn9 mus musculu
17	19	9.1	144	2	Q8BM42	Q8bm42 m mus muscu
18	19	9.1	148	2	Q75N02	Q75n02 homo sapien
19	19	9.1	508	1	CRB5_HUMAN	Q02930 homo sapien
20	9	4.3	850	2	Q9W0Z5	Q9w0z5 drosophila
21	8	3.8	141	2	Q8YWL5	Q8ywl5 anabaena sp
22	8	3.8	312	2	O62519	O62519 caenorhabdi
23	8	3.8	339	2	Q6DCE7	Q6dce7 xenopus lae
24	8	3.8	340	2	Q7V1Y2	Q7v1y2 prochloroco
25	8	3.8	384	2	Q73DB9	Q73db9 bacillus ce
26	8	3.8	393	2	Q7P753	Q7p753 fusobacteri
27	8	3.8	395	2	Q978Z6	Q978z6 thermoplasm
28	8	3.8	458	2	Q8I5J7	Q8i5j7 plasmodium
29	8	3.8	506	2	Q6ZJE1	Q6zje1 oryza sativ
30	8	3.8	575	2	Q6LG83	Q6lg83 photobacter
31	8	3.8	594	2	O73932	O73932 xenopus lae
32	8	3.8	599	2	Q93CJ2	Q93cj2 anabaena va
33	8	3.8	887	2	Q9MAQ3	Q9maq3 arabidopsis
34	8	3.8	887	2	Q9ZSB6	Q9zsb6 arabidopsis
35	8	3.8	910	2	Q93Y01	Q93y01 arabidopsis
36	8	3.8	928	2	Q9T0B6	Q9t0b6 arabidopsis
37	8	3.8	937	2	Q9T0B8	Q9t0b8 arabidopsis
38	8	3.8	951	2	Q8H557	Q8h557 oryza sativ
39	8	3.8	1028	2	Q9ZSB5	Q9zsb5 arabidopsis
40	8	3.8	1212	2	Q7RJZ3	Q7rjz3 plasmodium
41	8	3.8	1350	2	Q7T3T6	Q7t3t6 brachydanio
42	8	3.8	1476	2	Q9GRN0	Q9grn0 leishmania
43	7	3.3	55	2	Q653G0	Q653g0 oryza sativ
44	7	3.3	105	2	Q8C8W1	Q8c8w1 mus musculu
45	7	3.3	116	2	Q6Z8K8	Q6z8k8 oryza sativ

ALIGNMENTS

RESULT 1
Q8TAR1
ID Q8TAR1 PRELIMINARY; PRT; 209 AA.
AC Q8TAR1;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ATF2 protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC026175; AAH26175.1; -.
DR HSSP; P15336; 1BHI.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 1.
DR SMART; SM00355; ZnF_C2H2; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 1.
SQ SEQUENCE 209 AA; 23050 MW; A26AF07CA5D8D5E7 CRC64;

Query Match 100.0%; Score 209; DB 2; Length 209;

```
  Best Local Similarity   100.0%;  Pred. No. 1.8e-208;
  Matches  209;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MKFKLHVNSARQYKDLWNMSDDKPFLCTAPGCGQRFTNEDHLAVHKHKHEMTLKFGPARN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKFKLHVNSARQYKDLWNMSDDKPFLCTAPGCGQRFTNEDHLAVHKHKHEMTLKFGPARN 60

Qy         61 DSVIVADQTPTPTRFLKNCEEVGLFNELASPFENEFKKASEDDIKKMPLDLSPLATPIIR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DSVIVADQTPTPTRFLKNCEEVGLFNELASPFENEFKKASEDDIKKMPLDLSPLATPIIR 120

Qy        121 SKIEEPSVVETTHQDSPLPHPESTTSDEKEVPLAQTAQPTSAIVRPASLQVPNVLLTSSD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SKIEEPSVVETTHQDSPLPHPESTTSDEKEVPLAQTAQPTSAIVRPASLQVPNVLLTSSD 180

Qy        181 SSVIIQQAVPSPTSSTVITQAPSSNRPIV 209
              |||||||||||||||||||||||||||||
Db        181 SSVIIQQAVPSPTSSTVITQAPSSNRPIV 209


RESULT 2
Q8BN75
ID   Q8BN75      PRELIMINARY;      PRT;   234 AA.
AC   Q8BN75;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Mus musculus 0 day neonate eyeball cDNA, RIKEN full-length enriched
DE   library, clone:E130020I02 product:activating transcription factor 2,
DE   full insert sequence.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Eyeball;
RX   MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA   Carninci P., Hayashizaki Y.;
RT   "High-efficiency full-length cDNA cloning.";
RL   Meth. Enzymol. 303:19-44(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Eyeball;
RX   MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA   RIKEN FANTOM Consortium;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Eyeball;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Eyeball;
RX   MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA   Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA   Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT   "Normalization and subtraction of cap-trapper-selected cDNAs to
RT   prepare full-length cDNA libraries for rapid discovery of new genes.";
RL   Genome Res. 10:1617-1630(2000).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Eyeball;
RX   MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA   Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA   Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA   Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA   Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA   Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA   Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA   Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT   "RIKEN integrated sequence analysis (RISA) system-384-format
RT   sequencing pipeline with 384 multicapillary sequencer.";
RL   Genome Res. 10:1757-1771(2000).
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Eyeball;
RA   Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA   Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA   Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA   Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA   Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA   Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA   Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA   Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA   Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA   Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA   Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK087424; BAC39869.1; -.
DR   HSSP; P15336; 1BHI.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0003676; F:nucleic acid binding; IEA.
DR   GO; GO:0008270; F:zinc ion binding; IEA.
DR   InterPro; IPR007087; Znf_C2H2.
DR   Pfam; PF00096; zf-C2H2; 1.
DR   SMART; SM00355; ZnF_C2H2; 1.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR   PROSITE; PS50157; ZINC_FINGER_C2H2_2; 1.
SQ   SEQUENCE   234 AA;  25208 MW;  F06750FA9EB37A4D CRC64;

  Query Match             91.4%;  Score 191;  DB 2;  Length 234;
  Best Local Similarity   100.0%;  Pred. No. 1e-189;
  Matches  191;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         19 MSDDKPFLCTAPGCGQRFTNEDHLAVHKHKHEMTLKFGPARNDSVIVADQTPTPTRFLKN 78
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSDDKPFLCTAPGCGQRFTNEDHLAVHKHKHEMTLKFGPARNDSVIVADQTPTPTRFLKN 60

Qy         79 CEEVGLFNELASPFENEFKKASEDDIKKMPLDLSPLATPIIRSKIEEPSVVETTHQDSPL 138
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CEEVGLFNELASPFENEFKKASEDDIKKMPLDLSPLATPIIRSKIEEPSVVETTHQDSPL 120

Qy        139 PHPESTTSDEKEVPLAQTAQPTSAIVRPASLQVPNVLLTSSDSSVIIQQAVPSPTSSTVI 198
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 PHPESTTSDEKEVPLAQTAQPTSAIVRPASLQVPNVLLTSSDSSVIIQQAVPSPTSSTVI 180

Qy        199 TQAPSSNRPIV 209
              |||||||||||
Db        181 TQAPSSNRPIV 191


RESULT 3
Q8CBR9
ID   Q8CBR9      PRELIMINARY;      PRT;   420 AA.
AC   Q8CBR9;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Mus musculus adult male urinary bladder cDNA, RIKEN full-length
DE   enriched library, clone:9530046I22 product:activating transcription
DE   factor 2, full insert sequence.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Urinary bladder;
RX   MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA   Carninci P., Hayashizaki Y.;
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 8, 2005, 14:06:13 ; Search time 3168 Seconds
(without alignments)
2511.184 Million cell updates/sec

Title: AAH26175
Perfect score: 209
Sequence: 1 MKFKLHVNSARQYKDLWNMS.....PSPTSSTVITQAPSSNRPIV 209

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Word size: 1

Total number of hits satisfying chosen parameters: 68477535

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/AAH26175/runat_07072005_101848_20015/app_query.fasta_1.391
-DB=EST -QFMT=fastap -SUFFIX=oli.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=AAH26175_@CGN_1_1_4352_@runat_07072005_101848_20015 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	199	95.2	904	7	CR763073	CR763073 DKFZp469I
2	195	93.3	1606	3	AK087424	AK087424 Mus muscu
3	195	93.3	3623	3	AK035438	AK035438 Mus muscu
4	193	92.3	900	5	BU158697	BU158697 AGENCOURT
5	193	92.3	956	5	BU839869	BU839869 AGENCOURT
6	178	85.2	883	4	BI854668	BI854668 603381603
7	170	81.3	650	7	CN367893	CN367893 170005326
8	166	79.4	792	6	CA750476	CA750476 UI-M-FY0-
9	164	78.5	613	5	BU662818	BU662818 cl91g01.z
10	164	78.5	802	7	CR629564	CR629564 DKFZp469G
11	163	78.0	670	7	CN367891	CN367891 170004247
12	153	73.2	746	6	CD348287	CD348287 UI-M-FY0-
13	153	73.2	749	6	CD348288	CD348288 UI-M-FY0-
14	151	72.2	915	5	BQ715489	BQ715489 AGENCOURT
15	149	71.3	806	6	CB961368	CB961368 AGENCOURT
16	147	70.3	664	7	CF535749	CF535749 UI-M-GH0-
17	145	69.4	577	7	CN367889	CN367889 170004241
18	145	69.4	651	7	CN367890	CN367890 170005319
19	143	68.4	686	2	BE882603	BE882603 601507056
20	143	68.4	757	7	CK635513	CK635513 UI-M-HN0-
21	143	68.4	875	5	BQ891414	BQ891414 AGENCOURT
22	143	68.4	2106	3	AK051591	AK051591 Mus muscu
23	143	68.4	3579	3	AK041145	AK041145 Mus muscu
24	142	67.9	552	9	CL570065	CL570065 AS1007 Sa
25	138	66.0	780	7	CF742229	CF742229 UI-M-HB0-
26	135	64.6	763	7	CN367888	CN367888 170006000
27	135	64.6	958	5	BU503620	BU503620 AGENCOURT
28	134	64.1	417	8	CC249614	CC249614 XI426 Bay
29	133	63.6	818	7	CN455466	CN455466 UI-M-HN0-
30	132	63.2	742	7	CK634629	CK634629 UI-M-HN0-
31	132	63.2	758	6	CA510974	CA510974 UI-R-FJ0-
32	132	63.2	781	6	CA512317	CA512317 UI-R-FJ0-
33	132	63.2	2835	3	AY724488	AY724488 Rattus no
34	122	58.4	621	7	CN367892	CN367892 170005326
35	118	56.5	585	1	AI159526	AI159526 vz80c07.r
36	115	55.0	700	5	BU611698	BU611698 UI-M-FI0-
37	109	52.2	924	4	BI551997	BI551997 603195125
38	108	51.7	723	7	CN454827	CN454827 UI-M-HN0-
39	107	51.2	722	7	CF532094	CF532094 UI-M-FY0-
40	106	50.7	400	2	BE818106	BE818106 CM2-BN027
41	106	50.7	606	5	BQ037883	BQ037883 pgn1c.pk0
42	106	50.7	640	7	CN455108	CN455108 UI-M-HN0-
43	106	50.7	744	1	AJ453735	AJ453735 AJ453735
44	106	50.7	826	1	AJ454852	AJ454852 AJ454852
45	104	49.8	576	5	BP257160	BP257160 BP257160

ALIGNMENTS

RESULT 1
CR763073
LOCUS CR763073 904 bp mRNA linear EST 23-SEP-2004
DEFINITION DKFZp469I1824_r1 469 (synonym: pkid1) Pongo pygmaeus cDNA clone DKFZp469I1824 5', mRNA sequence.
ACCESSION CR763073
VERSION CR763073.1 GI:52600802
KEYWORDS EST.
SOURCE Pongo pygmaeus (orangutan)
ORGANISM Pongo pygmaeus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
REFERENCE 1 (bases 1 to 904)
AUTHORS Ansorge,W., Krieger,S., Regiert,T., Rittmueller,C., Schwager,B., Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.
TITLE Pongo pygmaeus mRNA (Ansorge,W., Krieger,S., Regiert,T., et al.)
JOURNAL Unpublished (2004)
COMMENT Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert. Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; rlin, Germany. Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp469I1824
Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.
FEATURES Location/Qualifiers
source 1. .904
/organism="Pongo pygmaeus"
/mol_type="mRNA"
/db_xref="taxon:9600"

```
                     /clone="DKFZp469I1824"
                     /tissue_type="kidney"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="469 (synonym: pkid1)"
                     /note="Vector: pSport1_Sfi; Site_1: SfilA; Site_2: SfilB"
ORIGIN

Alignment Scores:
Pred. No.:                4.07e-186      Length:          904
Score:                    199.00         Matches:         199
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              95.22%         Indels:          0
DB:                       7              Gaps:            0

AAH26175 (1-209) x CR763073 (1-904)

Qy          1 MetLysPheLysLeuHisValAsnSerAlaArgGlnTyrLysAspLeuTrpAsnMetSer 20
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        273 ATGAAATTCAAGTTACATGTGAATTCTGCCAGGCAATACAAGGACCTGTGGAATATGAGT 332

Qy         21 AspAspLysProPheLeuCysThrAlaProGlyCysGlyGlnArgPheThrAsnGluAsp 40
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        333 GATGACAAACCCTTTCTATGCACTGCGCCTGGATGTGGCCAGCGTTTTACCAACGAGGAT 392

Qy         41 HisLeuAlaValHisLysHisLysHisGluMetThrLeuLysPheGlyProAlaArgAsn 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        393 CATTTGGCTGTCCATAAACATAAACATGAGATGACACTGAAATTTGGTCCAGCACGTAAT 452

Qy         61 AspSerValIleValAlaAspGlnThrProThrProThrArgPheLeuLysAsnCysGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        453 GACAGTGTCATTGTGGCTGATCAGACCCCAACACCAACAAGATTCTTGAAAAACTGTGAA 512

Qy         81 GluValGlyLeuPheAsnGluLeuAlaSerProPheGluAsnGluPheLysLysAlaSer 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        513 GAAGTGGGTTTGTTTAATGAGTTGGCGAGTCCATTTGAGAATGAATTCAAGAAAGCTTCA 572

Qy        101 GluAspAspIleLysLysMetProLeuAspLeuSerProLeuAlaThrProIleIleArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        573 GAAGACGACATTAAAAAAATGCCTCTAGATTTATCCCCTCTTGCAACACCTATCATAAGA 632

Qy        121 SerLysIleGluGluProSerValValGluThrThrHisGlnAspSerProLeuProHis 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        633 AGCAAAATTGAGGAGCCTTCTGTTGTAGAAACGACTCACCAGGATAGTCCTTTACCTCAC 692

Qy        141 ProGluSerThrThrSerAspGluLysGluValProLeuAlaGlnThrAlaGlnProThr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        693 CCAGAGTCTACTACCAGTGATGAGAAGGAAGTACCACTGGCACAAACTGCACAGCCCACG 752

Qy        161 SerAlaIleValArgProAlaSerLeuGlnValProAsnValLeuLeuThrSerSerAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        753 TCAGCTATTGTTCGTCCAGCATCATTACAGGTTCCCAATGTGCTGCTTACAAGTTCTGAC 812

Qy        181 SerSerValIleIleGlnGlnAlaValProSerProThrSerSerThrValIleThr 199
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        813 TCAAGTGTAATTATTCAGCAGGCAGTACCTTCACCAACCTCAAGTACTGTAATCACC 869

RESULT 2
AK087424
LOCUS       AK087424               1606 bp    mRNA    linear   HTC 03-APR-2004
DEFINITION  Mus musculus 0 day neonate eyeball cDNA, RIKEN full-length enriched
            library, clone:E130020I02 product:activating transcription factor
            2, full insert sequence.
ACCESSION   AK087424
VERSION     AK087424.1  GI:26352477
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5
  AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6  (bases 1 to 1606)
  AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
            Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
            Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
            Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
            Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
            Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
            Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
            Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
            Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,
            URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
            Please visit our web site for further details.
            URL:http://genome.gsc.riken.jp/
            URL:http://fantom.gsc.riken.jp/.
FEATURES             Location/Qualifiers
     source          1. .1606
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:E130020I02"
                     /db_xref="taxon:10090"
                     /clone="E130020I02"
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 8, 2005, 12:19:52 ; Search time 26 Seconds
(without alignments)
773.435 Million cell updates/sec

Title: AAH26175
Perfect score: 209
Sequence: 1 MKFKLHVNSARQYKDLWNMS.....PSPTSSTVITQAPSSNRPIV 209

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	209	100.0	505	1	S05380	transcription fact
2	159	76.1	456	2	B42026	cyclic AMP respons
3	143	68.4	448	2	A42026	cAMP response elem
4	127	60.8	389	1	A39429	cAMP response elem
5	99	47.4	358	2	C42026	cyclic AMP respons
6	62	29.7	486	2	JC4028	activating transcr
7	55	26.3	313	2	A34785	DNA-binding protei
8	31	14.8	483	2	S12741	transcription fact
9	19	9.1	508	2	A45477	cAMP response elem
10	8	3.8	141	2	AH2004	hypothetical prote
11	8	3.8	242	2	T27999	hypothetical prote
12	8	3.8	887	2	C86453	CDS protein F9L11.
13	8	3.8	928	2	T04192	hypothetical prote
14	8	3.8	937	2	T04194	hypothetical prote
15	7	3.3	177	2	S74636	hypothetical prote
16	7	3.3	187	2	C90539	hypothetical prote
17	7	3.3	230	2	S35271	DNA-damage repair
18	7	3.3	237	2	A69255	polysaccharide ABC
19	7	3.3	255	2	S16151	myogenic factor Xm
20	7	3.3	268	2	D84849	hypothetical prote
21	7	3.3	287	2	T48548	hypothetical prote
22	7	3.3	355	2	T51414	CONSTANS-like 1 -
23	7	3.3	372	2	T32370	hypothetical prote
24	7	3.3	375	2	AD3432	A/G-specific adeni
25	7	3.3	400	2	H97201	probable flavoprot
26	7	3.3	425	2	H70456	nodulation competi
27	7	3.3	431	2	A89761	hypothetical prote
28	7	3.3	433	2	T43924	anthranilate synth
29	7	3.3	436	2	G69466	3-hydroxy-3-methyl
30	7	3.3	439	2	G97159	contains cell adhe
31	7	3.3	440	2	T44138	hypothetical prote
32	7	3.3	478	2	T08395	UTP-glucose glucos
33	7	3.3	513	2	A24309	genome polyprotein
34	7	3.3	520	2	T21462	hypothetical prote
35	7	3.3	522	1	GNWVR1	structural polypro
36	7	3.3	553	2	T45872	hypothetical prote
37	7	3.3	578	2	B90578	lipoprotein a [imp
38	7	3.3	612	2	C69840	conserved hypothet
39	7	3.3	726	2	S52141	ATPase - Erwinia a
40	7	3.3	769	2	F89870	serine proteinase
41	7	3.3	794	2	I58376	hypothetical prote
42	7	3.3	795	2	AC0129	probable surface a
43	7	3.3	859	2	S66827	aluminium resistan
44	7	3.3	963	2	T09478	ubiquitin thiolest
45	7	3.3	992	1	GNWVR3	structural polypro

ALIGNMENTS

RESULT 1
S05380
transcription factor ATF2 - human
N;Alternate names: activating transcription factor 2 (ATF-2); cAMP response element-bin
N;Contains: cAMP response element-binding protein HB16
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: S05380; A34776; B34223
R;Maekawa, T.; Sakura, H.; Kanei-Ishii, C.; Sudo, T.; Yoshimura, T.; Fujisawa, J.I.; Yo
EMBO J. 8, 2023-2028, 1989
A;Title: Leucine zipper structure of the protein CRE-BP1 binding to the cyclic AMP resp
A;Reference number: S05380; MUID:90005408; PMID:2529117
A;Accession: S05380
A;Molecule type: mRNA
A;Residues: 1-505 <MAE>
A;Cross-references: UNIPROT:P15336; EMBL:X15875; NID:g30214; PIDN:CAA33886.1; PID:g3021
R;Kara, C.J.; Liou, H.C.; Ivashkiv, L.B.; Glimcher, L.H.
Mol. Cell. Biol. 10, 1347-1357, 1990
A;Title: A cDNA for a human cyclic AMP response element-binding protein which is distin
A;Reference number: A34776; MUID:90205810; PMID:2320002
A;Accession: A34776
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 211-222,'N',224-505 <KAR>
A;Cross-references: GB:M31630; NID:g183787; PIDN:AAA35951.1; PID:g386762
R;Hai, T.; Liu, F.; Coukos, W.J.; Green, M.R.
Genes Dev. 3, 2083-2090, 1989
A;Title: Transcription factor ATF cDNA clones: an extensive family of leucine zipper pr
A;Reference number: A91622; MUID:90185187; PMID:2516827
A;Accession: B34223
A;Status: nucleic acid sequence not shown; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 107-357,'V',359-465 <HA2>
C;Genetics:
A;Gene: GDB:ATF2; CREB2; TREB7; CRE-BP1
A;Cross-references: GDB:128011; OMIM:123811
A;Map position: 2q32-2q32
C;Superfamily: cAMP response element-binding protein 1; fos/jun DNA-binding domain homo
C;Keywords: alternative splicing; DNA binding; nucleus; transcription regulation
F;347-387/Domain: fos/jun DNA-binding domain homology <FJD>

Query Match 100.0%; Score 209; DB 1; Length 505;
Best Local Similarity 100.0%; Pred. No. 8.4e-214;
Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MKFKLHVNSARQYKDLWNMSDDKPFLCTAPGCGQRFTNEDHLAVHKHKHEMTLKFGPARN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKFKLHVNSARQYKDLWNMSDDKPFLCTAPGCGQRFTNEDHLAVHKHKHEMTLKFGPARN 60

Qy         61 DSVIVADQTPTPTRFLKNCEEVGLFNELASPFENEFKKASEDDIKKMPLDLSPLATPIIR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DSVIVADQTPTPTRFLKNCEEVGLFNELASPFENEFKKASEDDIKKMPLDLSPLATPIIR 120
```

```
Qy        121 SKIEEPSVVETTHQDSPLPHPESTTSDEKEVPLAQTAQPTSAIVRPASLQVPNVLLTSSD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 SKIEEPSVVETTHQDSPLPHPESTTSDEKEVPLAQTAQPTSAIVRPASLQVPNVLLTSSD 180

Qy        181 SSVIIQQAVPSPTSSTVITQAPSSNRPIV 209
              |||||||||||||||||||||||||||||
Db        181 SSVIIQQAVPSPTSSTVITQAPSSNRPIV 209
```

RESULT 2
B42026
cyclic AMP response element DNA-binding protein isoform 1 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 26-Aug-1999
C;Accession: B42026
R;Georgopoulos, K.; Morgan, B.A.; Moore, D.D.
Mol. Cell. Biol. 12, 747-757, 1992
A;Title: Functionally distinct isoforms of the CRE-BP DNA-binding protein mediate activi
A;Reference number: A42026; MUID:92123199; PMID:1531087
A;Contents: EL4
A;Accession: B42026
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-456 <GEO>
A;Cross-references: GB:S76657; NID:g243428; PIDN:AAB21128.1; PID:g243429
A;Note: sequence extracted from NCBI backbone (NCBIN:76657, NCBIP:76658)
C;Superfamily: cAMP response element-binding protein 1; fos/jun DNA-binding domain homol
F;297-337/Domain: fos/jun DNA-binding domain homology <FJD>

  Query Match   76.1%; Score 159; DB 2; Length 456;
  Best Local Similarity   100.0%; Pred. No. 1.1e-160;
  Matches 159; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy         51 MTLKFGPARNDSVIVADQTPTPTRFLKNCEEVGLFNELASPFENEFKKASEDDIKKMPLD 110
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MTLKFGPARNDSVIVADQTPTPTRFLKNCEEVGLFNELASPFENEFKKASEDDIKKMPLD 60

Qy        111 LSPLATPIIRSKIEEPSVVETTHQDSPLPHPESTTSDEKEVPLAQTAQPTSAIVRPASLQ 170
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LSPLATPIIRSKIEEPSVVETTHQDSPLPHPESTTSDEKEVPLAQTAQPTSAIVRPASLQ 120

Qy        171 VPNVLLTSSDSSVIIQQAVPSPTSSTVITQAPSSNRPIV 209
              |||||||||||||||||||||||||||||||||||||||
Db        121 VPNVLLTSSDSSVIIQQAVPSPTSSTVITQAPSSNRPIV 159
```

RESULT 3
A42026
cAMP response element-binding protein 3 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 26-Aug-1999
C;Accession: A42026
R;Georgopoulos, K.; Morgan, B.A.; Moore, D.D.
Mol. Cell. Biol. 12, 747-757, 1992
A;Title: Functionally distinct isoforms of the CRE-BP DNA-binding protein mediate activi
A;Reference number: A42026; MUID:92123199; PMID:1531087
A;Contents: EL4
A;Accession: A42026
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-448 <GEO>
A;Cross-references: GB:S76655; NID:g243426; PIDN:AAB21127.1; PID:g243427
A;Note: sequence extracted from NCBI backbone (NCBIN:76655, NCBIP:76656)
C;Superfamily: cAMP response element-binding protein 1; fos/jun DNA-binding domain homol
F;289-329/Domain: fos/jun DNA-binding domain homology <FJD>

  Query Match   68.4%; Score 143; DB 2; Length 448;
  Best Local Similarity   100.0%; Pred. No. 1.1e-143;
  Matches 143; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy         67 DQTPTPTRFLKNCEEVGLFNELASPFENEFKKASEDDIKKMPLDLSPLATPIIRSKIEEP 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          9 DQTPTPTRFLKNCEEVGLFNELASPFENEFKKASEDDIKKMPLDLSPLATPIIRSKIEEP 68

Qy        127 SVVETTHQDSPLPHPESTTSDEKEVPLAQTAQPTSAIVRPASLQVPNVLLTSSDSSVIIQ 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         69 SVVETTHQDSPLPHPESTTSDEKEVPLAQTAQPTSAIVRPASLQVPNVLLTSSDSSVIIQ 128

Qy        187 QAVPSPTSSTVITQAPSSNRPIV 209
              |||||||||||||||||||||||
Db        129 QAVPSPTSSTVITQAPSSNRPIV 151
```

RESULT 4
A39429
cAMP response element-binding protein ATF2 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: A39429
R;Kageyama, R.; Sasai, Y.; Nakanishi, S.
J. Biol. Chem. 266, 15525-15531, 1991
A;Title: Molecular characterization of transcription factors that bind to the cAMP respo
A;Reference number: A39429; MUID:91332085; PMID:1714459
A;Accession: A39429
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-389 <KAG>
A;Cross-references: UNIPROT:Q00969; GB:M65148; NID:g206569; PIDN:AAA42013.1; PID:g206570
C;Superfamily: cAMP response element-binding protein 1; fos/jun DNA-binding domain homol
C;Keywords: DNA binding; nucleus; transcription regulation
F;231-271/Domain: fos/jun DNA-binding domain homology <FJD>

  Query Match   60.8%; Score 127; DB 1; Length 389;
  Best Local Similarity   100.0%; Pred. No. 1e-126;
  Matches 127; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy         19 MSDDKPFLCTAPGCGQRFTNEDHLAVHKHKHEMTLKFGPARNDSVIVADQTPTPTRFLKN 78
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSDDKPFLCTAPGCGQRFTNEDHLAVHKHKHEMTLKFGPARNDSVIVADQTPTPTRFLKN 60

Qy         79 CEEVGLFNELASPFENEFKKASEDDIKKMPLDLSPLATPIIRSKIEEPSVVETTHQDSPL 138
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CEEVGLFNELASPFENEFKKASEDDIKKMPLDLSPLATPIIRSKIEEPSVVETTHQDSPL 120

Qy        139 PHPESTT 145
              |||||||
Db        121 PHPESTT 127
```

RESULT 5
C42026
cyclic AMP response element DNA-binding protein isoform 2 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 26-Aug-1999
C;Accession: C42026
R;Georgopoulos, K.; Morgan, B.A.; Moore, D.D.
Mol. Cell. Biol. 12, 747-757, 1992
A;Title: Functionally distinct isoforms of the CRE-BP DNA-binding protein mediate activi
A;Reference number: A42026; MUID:92123199; PMID:1531087
A;Contents: EL4
A;Accession: C42026
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-358 <GEO>
A;Cross-references: GB:S76659; NID:g243430; PIDN:AAB21129.1; PID:g243431
A;Note: sequence extracted from NCBI backbone (NCBIN:76659, NCBIP:76660)
C;Superfamily: cAMP response element-binding protein 1; fos/jun DNA-binding domain homol
F;199-239/Domain: fos/jun DNA-binding domain homology <FJD>

  Query Match   47.4%; Score 99; DB 2; Length 358;
  Best Local Similarity   100.0%; Pred. No. 5.5e-97;
  Matches   99; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy         51 MTLKFGPARNDSVIVADQTPTPTRFLKNCEEVGLFNELASPFENEFKKASEDDIKKMPLD 110
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 8, 2005, 15:27:50 ; Search time 594 Seconds
(without alignments)
2209.037 Million cell updates/sec

Title: AAH26175
Perfect score: 209
Sequence: 1 MKFKLHVNSARQYKDLWNMS.....PSPTSSTVITQAPSSNRPIV 209

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 6330943 seqs, 3139157217 residues

Word size: 1

Total number of hits satisfying chosen parameters: 12650797

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/AAH26175/runat_07072005_101851_20099/app_query.fasta_1.391
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=oli.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=AAH26175_@CGN_1_1_741_@runat_07072005_101851_20099
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	57	27.3	501	9	US-09-833-790-215	Sequence 215, App
2	34	16.3	509	16	US-10-210-120-82	Sequence 82, Appl
3	34	16.3	509	22	US-10-909-035-82	Sequence 82, Appl
4	31	14.8	800	20	US-10-357-930-23141	Sequence 23141, A
5	31	14.8	800	20	US-10-357-930-29011	Sequence 29011, A
6	31	14.8	2755	20	US-10-357-930-22633	Sequence 22633, A
7	31	14.8	2755	20	US-10-357-930-28476	Sequence 28476, A
8	19	9.1	2345	20	US-10-723-860-7409	Sequence 7409, Ap
9	19	9.1	2637	11	US-09-968-007A-974	Sequence 974, App
10	19	9.1	2637	15	US-10-171-581-285	Sequence 285, App
11	19	9.1	2637	21	US-10-843-641A-7444	Sequence 7444, Ap
12	17	8.1	51	15	US-10-211-088-92	Sequence 92, Appl
13	16	7.7	51	15	US-10-211-088-80	Sequence 80, Appl
14	14	6.7	273	20	US-10-357-930-4968	Sequence 4968, Ap
15	13	6.2	397	20	US-10-357-930-14137	Sequence 14137, A
16	13	6.2	432	20	US-10-357-930-35258	Sequence 35258, A
17	13	6.2	432	20	US-10-357-930-44098	Sequence 44098, A
c 18	9	4.3	1121	19	US-10-767-701-10951	Sequence 10951, A
c 19	9	4.3	1264	18	US-10-424-599-98072	Sequence 98072, A
c 20	9	4.3	2775	9	US-09-738-626-2994	Sequence 2994, Ap
c 21	9	4.3	2809	21	US-10-494-672-95	Sequence 95, Appl
c 22	9	4.3	254366	10	US-09-822-871-3	Sequence 3, Appli
c 23	9	4.3	254366	18	US-10-673-885-3	Sequence 3, Appli
24	9	4.3	3309400	9	US-09-738-626-1	Sequence 1, Appli
c 25	8	3.8	139	9	US-09-728-444-136	Sequence 136, App
26	8	3.8	204	18	US-10-424-599-53079	Sequence 53079, A
c 27	8	3.8	275	19	US-10-674-124A-2643	Sequence 2643, Ap
c 28	8	3.8	290	9	US-09-294-093B-2862	Sequence 2862, Ap
29	8	3.8	296	9	US-09-867-701-8314	Sequence 8314, Ap
c 30	8	3.8	333	14	US-10-043-875-294	Sequence 294, App
c 31	8	3.8	333	14	US-10-043-875-307	Sequence 307, App
32	8	3.8	339	19	US-10-437-963-29382	Sequence 29382, A
33	8	3.8	378	20	US-10-425-115-47346	Sequence 47346, A
34	8	3.8	396	20	US-10-425-115-3184	Sequence 3184, Ap
35	8	3.8	406	19	US-10-437-963-87906	Sequence 87906, A
36	8	3.8	419	10	US-09-918-995-9121	Sequence 9121, Ap
c 37	8	3.8	431	9	US-09-764-847-1885	Sequence 1885, Ap
c 38	8	3.8	431	9	US-09-764-847-1886	Sequence 1886, Ap
c 39	8	3.8	431	14	US-10-092-154-1885	Sequence 1885, Ap
c 40	8	3.8	431	14	US-10-092-154-1886	Sequence 1886, Ap
c 41	8	3.8	440	9	US-09-764-847-449	Sequence 449, App
c 42	8	3.8	440	14	US-10-092-154-449	Sequence 449, App
43	8	3.8	441	20	US-10-425-115-161483	Sequence 161483,
44	8	3.8	463	9	US-09-954-531-1012	Sequence 1012, Ap
45	8	3.8	463	21	US-10-843-641A-2079	Sequence 2079, Ap

ALIGNMENTS

RESULT 1
US-09-833-790-215
; Sequence 215, Application US/09833790
; Patent No. US20020068288A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Wang, Tongtong
; APPLICANT: Secrist, Heather
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Indirias, Carol Y.
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.512
; CURRENT APPLICATION NUMBER: US/09/833,790
; CURRENT FILING DATE: 2001-04-11

```
;  NUMBER OF SEQ ID NOS: 440
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 215
;   LENGTH: 501
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(501)
;   OTHER INFORMATION: n = A,T,C or G
US-09-833-790-215

Alignment Scores:
Pred. No.:              1.15e-48        Length:         501
Score:                  57.00           Matches:        57
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            27.27%          Indels:         0
DB:                     9               Gaps:           0

AAH26175 (1-209) x US-09-833-790-215 (1-501)

Qy          1 MetLysPheLysLeuHisValAsnSerAlaArgGlnTyrLysAspLeuTrpAsnMetSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 ATGAAATTCAAGTTACATGTGAATTCTGCCAGGCAATACAAGGACCTGTGGAATATGAGT 326

Qy         21 AspAspLysProPheLeuCysThrAlaProGlyCysGlyGlnArgPheThrAsnGluAsp 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 GATGACAAACCCTTTCTATGTACTGCGCCTGGATGTGGCCAGCGTTTTACCAACGAGGAT 386

Qy         41 HisLeuAlaValHisLysHisLysHisGluMetThrLeuLysPheGlyPro 57
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 CATTTGGCTGTCCATAAACATAAACATGAGATGACACTGAAATTTGGTCCA 437

RESULT 2
US-10-210-120-82
; Sequence 82, Application US/10210120
; Publication No. US20030175736A1
; GENERAL INFORMATION:
;  APPLICANT: Chinnaiyan, Arul M.
;  APPLICANT:  Rubin, Mark A.
;  APPLICANT:  Sreekumar, Arun
;  TITLE OF INVENTION: Expression Profile of Prostate Cancer
;  FILE REFERENCE: UM-07221
;  CURRENT APPLICATION NUMBER: US/10/210,120
;  CURRENT FILING DATE:  2002-08-01
;  PRIOR APPLICATION NUMBER: US 60/309,581
;  PRIOR FILING DATE: 2001-08-02
;  PRIOR APPLICATION NUMBER: US 60/334,468
;  PRIOR FILING DATE: 2001-11-15
;  NUMBER OF SEQ ID NOS: 123
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 82
;   LENGTH: 509
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-210-120-82

Alignment Scores:
Pred. No.:              4.68e-25        Length:         509
Score:                  34.00           Matches:        34
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            16.27%          Indels:         0
DB:                     16              Gaps:           0

AAH26175 (1-209) x US-10-210-120-82 (1-509)

Qy          1 MetLysPheLysLeuHisValAsnSerAlaArgGlnTyrLysAspLeuTrpAsnMetSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAAATTCAAGTTACATGTGAATTCTGCCAGGCAATACAAGGACCTGTGGAATATGAGT 60

Qy         21 AspAspLysProPheLeuCysThrAlaProGlyCysGlyGln 34
              ||||||||||||||||||||||||||||||||||||||||||
Db         61 GATGACAAACCCTTTCTATGTACTGCGCCTGGATGTGGCCAG 102

RESULT 3
US-10-909-035-82
; Sequence 82, Application US/10909035
; Publication No. US20050136493A1
; GENERAL INFORMATION:
;  APPLICANT: Rubin, Mark A.
;  APPLICANT:  Chinnaiyan, Arul M.
;  APPLICANT:  Laxman, Bharathi
;  APPLICANT:  Sreekumar, Arun
;  TITLE OF INVENTION: AMACR Cancer Markers
;  FILE REFERENCE: UM-09098
;  CURRENT APPLICATION NUMBER: US/10/909,035
;  CURRENT FILING DATE:  2004-07-30
;  NUMBER OF SEQ ID NOS: 123
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 82
;   LENGTH: 509
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-909-035-82

Alignment Scores:
Pred. No.:              4.68e-25        Length:         509
Score:                  34.00           Matches:        34
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            16.27%          Indels:         0
DB:                     22              Gaps:           0

AAH26175 (1-209) x US-10-909-035-82 (1-509)

Qy          1 MetLysPheLysLeuHisValAsnSerAlaArgGlnTyrLysAspLeuTrpAsnMetSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAAATTCAAGTTACATGTGAATTCTGCCAGGCAATACAAGGACCTGTGGAATATGAGT 60

Qy         21 AspAspLysProPheLeuCysThrAlaProGlyCysGlyGln 34
              ||||||||||||||||||||||||||||||||||||||||||
Db         61 GATGACAAACCCTTTCTATGTACTGCGCCTGGATGTGGCCAG 102

RESULT 4
US-10-357-930-23141
; Sequence 23141, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
;  APPLICANT: Schlegel. Robert
;  APPLICANT:  Endege, Wilson
;  APPLICANT:  Monahan, John
;  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
;  TITLE OF INVENTION:  IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
;  TITLE OF INVENTION:  HUMAN PROSTATE CANCER
;  FILE REFERENCE: MRI-007BCN
;  CURRENT APPLICATION NUMBER: US/10/357,930
;  CURRENT FILING DATE:  2003-02-04
;  PRIOR APPLICATION NUMBER: 09/785,276
;  PRIOR FILING DATE: 2003-02-16
;  PRIOR APPLICATION NUMBER: 60/183,319
;  PRIOR FILING DATE: 2000-02-17
;  PRIOR APPLICATION NUMBER: 60/189,862
;  PRIOR FILING DATE: 2000-03-16
;  PRIOR APPLICATION NUMBER: 60/207,454
;  PRIOR FILING DATE: 2000-05-25
;  PRIOR APPLICATION NUMBER: 60/211,314
;  PRIOR FILING DATE: 2000-06-09
;  PRIOR APPLICATION NUMBER: 60/219,007
;  PRIOR FILING DATE: 2000-07-18
;  PRIOR APPLICATION NUMBER: 60/255,281
;  PRIOR FILING DATE: 2000-12-13
;  NUMBER OF SEQ ID NOS: 62232
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 23141
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 8, 2005, 14:07:03 ; Search time 183 Seconds
(without alignments)
1868.753 Million cell updates/sec

Title: AAH26175
Perfect score: 209
Sequence: 1 MKFKLHVNSARQYKDLWNMS.....PSPTSSTVITQAPSSNRPIV 209

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Word size: 1

Total number of hits satisfying chosen parameters: 2400006

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=x1h
-Q=/cgn2_1/USPTO_spool/AAH26175/runat_07072005_101849_20027/app_query.fasta_1.391
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=oli.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=AAH26175_@CGN_1_1_177_@runat_07072005_101849_20027 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	209	100.0	1621	4	US-09-949-016-5410	Sequence 5410, Ap
	2	60	28.7	65966	4	US-09-949-016-17152	Sequence 17152, A
	3	31	14.8	229	4	US-09-513-999C-2946	Sequence 2946, Ap
c	4	31	14.8	601	4	US-09-949-016-189073	Sequence 189073,
c	5	24	11.5	601	4	US-09-949-016-189064	Sequence 189064,
c	6	11	5.3	601	4	US-09-949-016-189057	Sequence 189057,
	7	9	4.3	3346	4	US-09-902-540-548	Sequence 548, App
c	8	9	4.3	157822	4	US-09-949-016-16723	Sequence 16723, A
c	9	9	4.3	254366	4	US-09-822-871-3	Sequence 3, Appli
c	10	8	3.8	158	4	US-09-513-999C-26841	Sequence 26841, A
c	11	8	3.8	462	4	US-09-248-796A-8283	Sequence 8283, Ap
	12	8	3.8	601	4	US-09-949-016-21288	Sequence 21288, A
	13	8	3.8	601	4	US-09-949-016-56382	Sequence 56382, A
	14	8	3.8	601	4	US-09-949-016-56383	Sequence 56383, A
c	15	8	3.8	601	4	US-09-949-016-83959	Sequence 83959, A
c	16	8	3.8	601	4	US-09-949-016-83960	Sequence 83960, A
	17	8	3.8	601	4	US-09-949-016-137071	Sequence 137071,
	18	8	3.8	601	4	US-09-949-016-137072	Sequence 137072,
	19	8	3.8	601	4	US-09-949-016-137073	Sequence 137073,
c	20	8	3.8	601	4	US-09-949-016-143248	Sequence 143248,
c	21	8	3.8	601	4	US-09-949-016-143419	Sequence 143419,
	22	8	3.8	601	4	US-09-949-016-163200	Sequence 163200,
	23	8	3.8	601	4	US-09-949-016-179127	Sequence 179127,
	24	8	3.8	601	4	US-09-949-016-191387	Sequence 191387,
c	25	8	3.8	1001	3	US-09-641-638-205	Sequence 205, App
c	26	8	3.8	1001	3	US-09-641-638-206	Sequence 206, App
c	27	8	3.8	1001	4	US-10-170-097-205	Sequence 205, App
c	28	8	3.8	1001	4	US-10-170-097-206	Sequence 206, App
c	29	8	3.8	1242	4	US-09-489-039A-1286	Sequence 1286, Ap
c	30	8	3.8	1365	4	US-09-857-556A-17	Sequence 17, Appl
	31	8	3.8	1488	4	US-09-949-016-3851	Sequence 3851, Ap
c	32	8	3.8	2689	2	US-08-876-546A-15	Sequence 15, Appl
c	33	8	3.8	2689	3	US-09-412-252-15	Sequence 15, Appl
	34	8	3.8	3162	4	US-09-328-352-932	Sequence 932, App
	35	8	3.8	5069	4	US-09-902-540-591	Sequence 591, App
c	36	8	3.8	5207	4	US-09-858-664A-1	Sequence 1, Appli
c	37	8	3.8	5207	4	US-10-274-978-1	Sequence 1, Appli
c	38	8	3.8	5207	4	US-10-274-978-3	Sequence 3, Appli
c	39	8	3.8	5207	4	US-10-697-263-1	Sequence 1, Appli
c	40	8	3.8	5207	4	US-10-697-263-3	Sequence 3, Appli
c	41	8	3.8	9663	4	US-09-869-588-22	Sequence 22, Appl
c	42	8	3.8	9726	4	US-09-869-588-21	Sequence 21, Appl
	43	8	3.8	9813	4	US-09-949-016-13582	Sequence 13582, A
	44	8	3.8	12047	2	US-09-022-461-1	Sequence 1, Appli
	45	8	3.8	12047	3	US-09-033-556-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-949-016-5410
; Sequence 5410, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5410
; LENGTH: 1621
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-5410

Alignment Scores:

Pred. No.:	1.17e-207	Length:	1621
Score:	209.00	Matches:	209
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

AAH26175 (1-209) x US-09-949-016-5410 (1-1621)

Qy 1 MetLysPheLysLeuHisValAsnSerAlaArgGlnTyrLysAspLeuTrpAsnMetSer 20

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          24 ATGAAATTCAAGTTACATGTGAATTCTGCCAGGCAATACAAGGACCTGTGGAATATGAGT 83

Qy          21 AspAspLysProPheLeuCysThrAlaProGlyCysGlyGlnArgPheThrAsnGluAsp 40
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          84 GATGACAAACCCTTTCTATGTACTGCGCCTGGATGTGGCCAGCGTTTTACCAACGAGGAT 143

Qy          41 HisLeuAlaValHisLysHisLysHisGluMetThrLeuLysPheGlyProAlaArgAsn 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         144 CATTTGGCTGTCCATAAACATAAACATGAGATGACACTGAAATTTGGTCCAGCACGTAAT 203

Qy          61 AspSerValIleValAlaAspGlnThrProThrProThrArgPheLeuLysAsnCysGlu 80
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         204 GACAGTGTCATTGTGGCTGATCAGACCCCAACACCAACAAGATTCTTGAAAAACTGTGAA 263

Qy          81 GluValGlyLeuPheAsnGluLeuAlaSerProPheGluAsnGluPheLysLysAlaSer 100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         264 GAAGTGGGTTTGTTTAATGAGTTGGCGAGTCCATTTGAGAATGAATTCAAGAAAGCTTCA 323

Qy         101 GluAspAspIleLysLysMetProLeuAspLeuSerProLeuAlaThrProIleIleArg 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         324 GAAGATGACATTAAAAAAATGCCTCTAGATTTATCCCCTCTTGCAACACCTATCATAAGA 383

Qy         121 SerLysIleGluGluProSerValValGluThrThrHisGlnAspSerProLeuProHis 140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         384 AGCAAAATTGAGGAGCCTTCTGTTGTAGAAACAACTCACCAGGATAGTCCTTTACCTCAC 443

Qy         141 ProGluSerThrThrSerAspGluLysGluValProLeuAlaGlnThrAlaGlnProThr 160
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         444 CCAGAGTCTACTACCAGTGATGAGAAGGAAGTACCATTGGCACAAACTGCACAGCCCACA 503

Qy         161 SerAlaIleValArgProAlaSerLeuGlnValProAsnValLeuLeuThrSerSerAsp 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         504 TCAGCTATTGTTCGTCCAGCATCATTACAGGTTCCCAATGTGCTGCTTACAAGTTCTGAC 563

Qy         181 SerSerValIleIleGlnGlnAlaValProSerProThrSerSerThrValIleThrGln 200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         564 TCAAGTGTAATTATTCAGCAGGCAGTACCTTCACCAACCTCAAGTACTGTAATCACCCAG 623

Qy         201 AlaProSerSerAsnArgProIleVal 209
                ||||||||||||||||||||||||||
Db         624 GCACCATCCTCTAACAGGCCAATTGTC 650

RESULT 2
US-09-949-016-17152
; Sequence 17152, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17152
;   LENGTH: 65966
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-17152

Alignment Scores:
Pred. No.:              4.44e-51        Length:         65966
Score:                  60.00           Matches:        60
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            28.71%          Indels:         0
DB:                     4               Gaps:           0

AAH26175 (1-209) x US-09-949-016-17152 (1-65966)

Qy         150 GluValProLeuAlaGlnThrAlaGlnProThrSerAlaIleValArgProAlaSerLeu 169
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       23601 GAAGTACCATTGGCACAAACTGCACAGCCCACATCAGCTATTGTTCGTCCAGCATCATTA 23660

Qy         170 GlnValProAsnValLeuLeuThrSerSerAspSerSerValIleIleGlnGlnAlaVal 189
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       23661 CAGGTTCCCAATGTGCTGCTTACAAGTTCTGACTCAAGTGTAATTATTCAGCAGGCAGTA 23720

Qy         190 ProSerProThrSerSerThrValIleThrGlnAlaProSerSerAsnArgProIleVal 209
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       23721 CCTTCACCAACCTCAAGTACTGTAATCACCCAGGCACCATCCTCTAACAGGCCAATTGTG 23780

RESULT 3
US-09-513-999C-2946
; Sequence 2946, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 2946
;   LENGTH: 229
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 72..227
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 87
;   OTHER INFORMATION: y=c or t
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: 6
;   OTHER INFORMATION: Xaa=Pro or Ser
US-09-513-999C-2946

Alignment Scores:
Pred. No.:              2.37e-23        Length:         229
Score:                  31.00           Matches:        31
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            14.83%          Indels:         0
DB:                     4               Gaps:           0

AAH26175 (1-209) x US-09-513-999C-2946 (1-229)

Qy          29 AlaProGlyCysGlyGlnArgPheThrAsnGluAspHisLeuAlaValHisLysHisLys 48
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         102 GCCCCGGGCTGTGGACAGAGATTTACAAACGAGGACCACCTGGCAGTTCATAAACACAAG 161

Qy          49 HisGluMetThrLeuLysPheGlyProAlaArg 59
                |||||||||||||||||||||||||||||||
Db         162 CATGAGATGACATTGAAATTTGGCCCAGCCCGA 194

RESULT 4
US-09-949-016-189073/c
; Sequence 189073, Application US/09949016
; Patent No. 6812339
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 8, 2005, 14:03:03 ; Search time 530 Seconds
(without alignments)
2334.389 Million cell updates/sec

Title: AAH26175
Perfect score: 209
Sequence: 1 MKFKLHVNSARQYKDLWNMS.....PSPTSSTVITQAPSSNRPIV 209

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Word size: 1

Total number of hits satisfying chosen parameters: 8776198

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/AAH26175/runat_07072005_101846_19958/app_query.fasta_1.391
-DB=N_Geneseq_16Dec04 -QFMT=fastap -SUFFIX=oli.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=AAH26175_@CGN_1_1_644_@runat_07072005_101846_19958 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	209	100.0	1647	10	ADK67038	Adk67038 Gene #128
2	209	100.0	1647	10	ADK66978	Adk66978 Gene #68
3	144	68.9	3507	13	ADR08447	Adr08447 Full leng
4	132	63.2	1852	10	ADB53173	Adb53173 Primary r
5	58	27.8	3217	6	ABI99202	Abi99202 Mouse isc
6	57	27.3	501	6	AAS61674	Aas61674 Lung smal
7	34	16.3	509	10	ADD18510	Add18510 Human pro
8	31	14.8	229	3	AAC02948	Aac02948 Human sec
9	31	14.8	800	5	ABV28993	Abv28993 Human pro
10	31	14.8	800	5	ABV23152	Abv23152 Human pro
11	31	14.8	1239	12	ADP80924	Adp80924 Mouse C13
12	31	14.8	2755	5	ABV28463	Abv28463 Human pro
13	31	14.8	2755	5	ABV22640	Abv22640 Human pro
14	19	9.1	2319	10	ADC18732	Adc18732 DNA of hu
15	19	9.1	2345	12	ADQ24589	Adq24589 Human sof
16	19	9.1	2637	6	ABL69107	Abl69107 Kidney ca
17	19	9.1	2637	6	ABK83799	Abk83799 Human cDN
18	19	9.1	2637	10	ADH29017	Adh29017 Human chr
19	17	8.1	51	10	ADC22243	Adc22243 Protein b
20	16	7.7	51	10	ADC22231	Adc22231 Protein b
21	14	6.7	273	5	ABV04977	Abv04977 Human pro
22	13	6.2	397	5	ABV14146	Abv14146 Human pro
23	13	6.2	432	5	ABV44079	Abv44079 Human pro
24	13	6.2	432	5	ABV35240	Abv35240 Human pro
25	9	4.3	31	4	AAI30980	Aai30980 Human sin
c 26	9	4.3	527	4	AAK78581	Aak78581 Human imm
c 27	9	4.3	527	4	AAK78582	Aak78582 Human imm
28	9	4.3	1080	4	ABL21331	Abl21331 Drosophil
c 29	9	4.3	2775	5	AAH67959	Aah67959 C glutami
c 30	9	4.3	2775	8	ACA02052	Aca02052 C. glutam
c 31	9	4.3	2809	11	ADL65738	Adl65738 C. glutam
32	9	4.3	3199	4	ABL21330	Abl21330 Drosophil
33	9	4.3	9397	4	AAK78065	Aak78065 Human imm
c 34	9	4.3	254366	8	ABZ23704	Abz23704 Human pho
35	9	4.3	349980	5	AAH68533	Aah68533 C glutami
c 36	8	3.8	139	8	ABX90792	Abx90792 Murine ge
c 37	8	3.8	158	3	AAC22766	Aac22766 Human sec
c 38	8	3.8	290	6	ABL73488	Abl73488 Corn tass
39	8	3.8	296	6	ABL85336	Abl85336 Human ova
c 40	8	3.8	333	6	ABZ34052	Abz34052 HIV-1 rev
c 41	8	3.8	333	6	ABZ34065	Abz34065 HIV-1 rev
c 42	8	3.8	410	4	AAI86431	Aai86431 Human pol
43	8	3.8	419	9	ACH21909	Ach21909 Human adu
c 44	8	3.8	431	4	ABK42999	Abk42999 Genomic s
c 45	8	3.8	431	4	ABK42998	Abk42998 Genomic s

ALIGNMENTS

RESULT 1
ADK67038
ID ADK67038 standard; DNA; 1647 BP.
XX
AC ADK67038;
XX
DT 06-MAY-2004 (first entry)
XX
DE Gene #128 for inhibitory RNA to manipulate stem cell phenotype.
XX
KW ds; gene; inhibitory RNA; RNAi; Notch signaling target gene; phenotype;
KW pluripotent stem cell.
XX
OS Homo sapiens.
XX
PN WO2003068961-A2.
XX
PD 21-AUG-2003.
XX
PF 12-FEB-2003; 2003WO-GB000579.
XX
PR 13-FEB-2002; 2002GB-00003359.
PR 13-FEB-2002; 2002GB-00003387.
XX
PA (AXOR-) AXORDIA LTD.
XX
PI Andrews P, Walsh J, Gokhale P;
XX
DR WPI; 2003-697528/66.

XX
PT New inhibitory RNA molecule having double stranded RNA molecules, useful
PT for manipulating the phenotype of stem cells, preferably pluripotential
PT stem cells.
XX
PS Disclosure; SEQ ID NO 128; 157pp; English.
XX
CC The invention relates to an inhibitory RNA (RNAi) molecule derived from a
CC nucleic acid molecule comprising a defined nucleic acid sequences given
CC in the specification or a sequence which hybridizes to the sequences and
CC encodes a Notch signaling target gene or which is a degenerate as a
CC result of the genetic code of the sequences. The methods and compositions
CC of the present invention are useful for manipulating the phenotype of
CC stem cells, preferably pluripotent stem cells. This sequence corresponds
CC to one of the nucleic acid molecules of the invention.
XX
SQ Sequence 1647 BP; 504 A; 396 C; 341 G; 406 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.01e-199	Length:	1647
Score:	209.00	Matches:	209
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0

AAH26175 (1-209) x ADK67038 (1-1647)

```
Qy          1 MetLysPheLysLeuHisValAsnSerAlaArgGlnTyrLysAspLeuTrpAsnMetSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         27 ATGAAATTCAAGTTACATGTGAATTCTGCCAGGCAATACAAGGACCTGTGGAATATGAGT 86

Qy         21 AspAspLysProPheLeuCysThrAlaProGlyCysGlyGlnArgPheThrAsnGluAsp 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         87 GATGACAAACCCTTTCTATGTACTGCGCCTGGATGTGGCCAGCGTTTTACCAACGAGGAT 146

Qy         41 HisLeuAlaValHisLysHisLysHisGluMetThrLeuLysPheGlyProAlaArgAsn 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 CATTTGGCTGTCCATAAACATAAACATGAGATGACACTGAAATTTGGTCCAGCACGTAAT 206

Qy         61 AspSerValIleValAlaAspGlnThrProThrProThrArgPheLeuLysAsnCysGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 GACAGTGTCATTGTGGCTGATCAGACCCCAACACCAACAAGATTCTTGAAAAACTGTGAA 266

Qy         81 GluValGlyLeuPheAsnGluLeuAlaSerProPheGluAsnGluPheLysLysAlaSer 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 GAAGTGGGTTTGTTTAATGAGTTGGCGAGTCCATTTGAGAATGAATTCAAGAAAGCTTCA 326

Qy        101 GluAspAspIleLysLysMetProLeuAspLeuSerProLeuAlaThrProIleIleArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 GAAGATGACATTAAAAAAATGCCTCTAGATTTATCCCCTCTTGCAACACCTATCATAAGA 386

Qy        121 SerLysIleGluGluProSerValValGluThrThrHisGlnAspSerProLeuProHis 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 AGCAAAATTGAGGAGCCTTCTGTTGTAGAAACAACTCACCAGGATAGTCCTTTACCTCAC 446

Qy        141 ProGluSerThrThrSerAspGluLysGluValProLeuAlaGlnThrAlaGlnProThr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 CCAGAGTCTACTACCAGTGATGAGAAGGAAGTACCATTGGCACAAACTGCACAGCCCACA 506

Qy        161 SerAlaIleValArgProAlaSerLeuGlnValProAsnValLeuLeuThrSerSerAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 TCAGCTATTGTTCGTCCAGCATCATTACAGGTTCCCAATGTGCTGCTTACAAGTTCTGAC 566

Qy        181 SerSerValIleIleGlnGlnAlaValProSerProThrSerSerThrValIleThrGln 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 TCAAGTGTAATTATTCAGCAGGCAGTACCTTCACCAACCTCAAGTACTGTAATCACCCAG 626

Qy        201 AlaProSerSerAsnArgProIleVal 209
              ||||||||||||||||||||||||||
Db        627 GCACCATCCTCTAACAGGCCAATTGTC 653
```

RESULT 2

ADK66978
ID ADK66978 standard; DNA; 1647 BP.
XX
AC ADK66978;
XX
DT 06-MAY-2004 (first entry)
XX
DE Gene #68 for inhibitory RNA to manipulate stem cell phenotype.
XX
KW ds; gene; inhibitory RNA; RNAi; Notch signaling target gene; phenotype;
KW pluripotent stem cell.
XX
OS Homo sapiens.
XX
PN WO2003068961-A2.
XX
PD 21-AUG-2003.
XX
PF 12-FEB-2003; 2003WO-GB000579.
XX
PR 13-FEB-2002; 2002GB-00003359.
PR 13-FEB-2002; 2002GB-00003387.
XX
PA (AXOR-) AXORDIA LTD.
XX
PI Andrews P, Walsh J, Gokhale P;
XX
DR WPI; 2003-697528/66.
XX
PT New inhibitory RNA molecule having double stranded RNA molecules, useful
PT for manipulating the phenotype of stem cells, preferably pluripotential
PT stem cells.
XX
PS Disclosure; SEQ ID NO 68; 157pp; English.
XX
CC The invention relates to an inhibitory RNA (RNAi) molecule derived from a
CC nucleic acid molecule comprising a defined nucleic acid sequences given
CC in the specification or a sequence which hybridizes to the sequences and
CC encodes a Notch signaling target gene or which is a degenerate as a
CC result of the genetic code of the sequences. The methods and compositions
CC of the present invention are useful for manipulating the phenotype of
CC stem cells, preferably pluripotent stem cells. This sequence corresponds
CC to one of the nucleic acid molecules of the invention.
XX
SQ Sequence 1647 BP; 504 A; 396 C; 341 G; 406 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.01e-199	Length:	1647
Score:	209.00	Matches:	209
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0

AAH26175 (1-209) x ADK66978 (1-1647)

```
Qy          1 MetLysPheLysLeuHisValAsnSerAlaArgGlnTyrLysAspLeuTrpAsnMetSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         27 ATGAAATTCAAGTTACATGTGAATTCTGCCAGGCAATACAAGGACCTGTGGAATATGAGT 86

Qy         21 AspAspLysProPheLeuCysThrAlaProGlyCysGlyGlnArgPheThrAsnGluAsp 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         87 GATGACAAACCCTTTCTATGTACTGCGCCTGGATGTGGCCAGCGTTTTACCAACGAGGAT 146

Qy         41 HisLeuAlaValHisLysHisLysHisGluMetThrLeuLysPheGlyProAlaArgAsn 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 CATTTGGCTGTCCATAAACATAAACATGAGATGACACTGAAATTTGGTCCAGCACGTAAT 206

Qy         61 AspSerValIleValAlaAspGlnThrProThrProThrArgPheLeuLysAsnCysGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 GACAGTGTCATTGTGGCTGATCAGACCCCAACACCAACAAGATTCTTGAAAAACTGTGAA 266

Qy         81 GluValGlyLeuPheAsnGluLeuAlaSerProPheGluAsnGluPheLysLysAlaSer 100
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 8, 2005, 14:05:28 ; Search time 4144 Seconds
(without alignments)
2443.808 Million cell updates/sec

Title: AAH26175
Perfect score: 209
Sequence: 1 MKFKLHVNSARQYKDLWNMS.....PSPTSSTVITQAPSSNRPIV 209

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Word size: 1

Total number of hits satisfying chosen parameters: 9408497

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/AAH26175/runat_07072005_101847_19991/app_query.fasta_1.391
-DB=GenEmbl -QFMT=fastap -SUFFIX=oli.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=AAH26175_@CGN_1_1_4200_@runat_07072005_101847_19991 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	209	100.0	1370	9	BC026175	BC026175 Homo sapi
2	209	100.0	1647	6	AX821940	AX821940 Sequence
3	209	100.0	1647	6	AX822000	AX822000 Sequence
4	209	100.0	1647	9	HSCREBP1	X15875 Human mRNA
5	195	93.3	3957	10	BC042210	BC042210 Mus muscu
6	191	91.4	1464	10	AF483482	AF483482 Mus muscu
7	191	91.4	1464	10	AF483483	AF483483 Mus muscu
8	183	87.6	1596	10	S76657	S76657 cyclic AMP
9	144	68.9	3507	6	CQ851484	CQ851484 Sequence
10	144	68.9	3507	9	AK128731	AK128731 Homo sapi
11	143	68.4	1848	10	S76655	S76655 cyclic AMP
12	143	68.4	3983	9	HSM808617	BX648469 Homo sapi
13	143	68.4	4130	10	BC079883	BC079883 Mus muscu
14	135	64.6	3909	10	BC082596	BC082596 Mus muscu
15	132	63.2	1209	10	RATRATF2	M65148 Rat RATF2 m
16	132	63.2	1852	10	RNU38938	U38938 Rattus norv
17	123	58.9	1302	10	S76659	S76659 cyclic AMP
18	106	50.7	1812	5	GGA17724	Y17724 Gallus gall
19	96	45.9	4851	12	AF050498	AF050498 Fusion tr
20	62	29.7	1520	5	XLU16158	U16158 Xenopus lae
c 21	60	28.7	176033	2	AL161794	AL161794 Homo sapi
c 22	60	28.7	178342	9	AC007435	AC007435 Homo sapi
c 23	60	28.7	207639	10	AL844581	AL844581 Mouse DNA
24	58	27.8	3217	6	AX305250	AX305250 Sequence
25	58	27.8	3217	10	MUSTCRATF	M77167 Mouse T-cel
c 26	58	27.8	166766	2	AC119682	AC119682 Rattus no
c 27	58	27.8	173793	2	AC129355	AC129355 Rattus no
28	57	27.3	501	6	AX321198	AX321198 Sequence
29	55	26.3	957	10	MUSMXBP	M31629 Mouse mXBP
30	34	16.3	508	9	AY029364	AY029364 Homo sapi
c 31	33	15.8	11797	9	AC074291	AC074291 Homo sapi
32	31	14.8	229	6	AX887083	AX887083 Sequence
33	31	14.8	229	6	BD026693	BD026693 Sequence
34	31	14.8	800	6	CQ491274	CQ491274 Sequence
35	31	14.8	800	6	CQ497144	CQ497144 Sequence
36	31	14.8	1227	6	CQ720976	CQ720976 Sequence
37	31	14.8	1485	9	HSATFA1	X57197 H.sapiens m
38	31	14.8	2755	6	CQ490766	CQ490766 Sequence
39	31	14.8	2755	6	CQ496609	CQ496609 Sequence
40	31	14.8	2758	9	HSATFA	X52943 Human mRNA
41	31	14.8	2867	10	BC026483	BC026483 Mus muscu
42	29	13.9	787	9	BC042363	BC042363 Homo sapi
43	26	12.4	172558	9	AC023509	AC023509 Homo sapi
44	26	12.4	223132	10	AC123870	AC123870 Mus muscu
c 45	26	12.4	243444	2	AC133748	AC133748 Rattus no

ALIGNMENTS

RESULT 1
BC026175
LOCUS BC026175 1370 bp mRNA linear PRI 16-SEP-2003
DEFINITION Homo sapiens activating transcription factor 2, mRNA (cDNA clone IMAGE:4308091), complete cds.
ACCESSION BC026175
VERSION BC026175.1 GI:20072896
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1370)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,

```
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1370)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-APR-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
            cDNA Library Preparation: CLONTECH Laboratories, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Institute for Systems Biology
            http://www.systemsbiology.org
            contact: amadan@systemsbiology.org
            Anup Madan, Jessica Fahey, Erin Helton, Mark Ketteman, Anuradha
            Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 38 Row: c Column: 20
            This clone has the following problem: The cds is short compared to
            the longest cds in the locus.
FEATURES             Location/Qualifiers
     source          1. .1370
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4308091"
                     /tissue_type="Skin, melanotic melanoma, high MDR."
                     /clone_lib="NIH_MGC_62"
                     /lab_host="DH10B"
                     /note="Vector: pDNR-LIB"
     gene            1. .1370
                     /gene="ATF2"
                     /note="synonyms: CRE-BP1, TREB7, HB16"
                     /db_xref="LocusID:1386"
                     /db_xref="MIM:123811"
     CDS             267. .896
                     /gene="ATF2"
                     /codon_start=1
                     /product="ATF2 protein"
                     /protein_id="AAH26175.1"
                     /db_xref="GI:20072897"
                     /db_xref="LocusID:1386"
                     /db_xref="MIM:123811"
                     /translation="MKFKLHVNSARQYKDLWNMSDDKPFLCTAPGCGQRFTNEDHLAV
                     HKHKHEMTLKFGPARNDSVIVADQTPTPTRFLKNCEEVGLFNELASPFENEFKKASED
                     DIKKMPLDLSPLATPIIRSKIEEPSVVETTHQDSPLPHPESTTSDEKEVPLAQTAQPT
                     SAIVRPASLQVPNVLLTSSDSSVIIQQAVPSPTSSTVITQAPSSNRPIV"
ORIGIN

Alignment Scores:
Pred. No.:              1.96e-209       Length:         1370
Score:                  209.00          Matches:        209
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     9               Gaps:           0

AAH26175 (1-209) x BC026175 (1-1370)

Qy          1 MetLysPheLysLeuHisValAsnSerAlaArgGlnTyrLysAspLeuTrpAsnMetSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        267 ATGAAATTCAAGTTACATGTGAATTCTGCCAGGCAATACAAGGACCTGTGGAATATGAGT 326

Qy         21 AspAspLysProPheLeuCysThrAlaProGlyCysGlyGlnArgPheThrAsnGluAsp 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        327 GATGACAAACCCTTTCTATGTACTGCGCCTGGATGTGGCCAGCGTTTTACCAACGAGGAT 386

Qy         41 HisLeuAlaValHisLysHisLysHisGluMetThrLeuLysPheGlyProAlaArgAsn 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        387 CATTTGGCTGTCCATAAACATAAACATGAGATGACACTGAAATTTGGTCCAGCACGTAAT 446

Qy         61 AspSerValIleValAlaAspGlnThrProThrProThrArgPheLeuLysAsnCysGlu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        447 GACAGTGTCATTGTGGCTGATCAGACCCCAACACCAACAAGATTCTTGAAAAACTGTGAA 506

Qy         81 GluValGlyLeuPheAsnGluLeuAlaSerProPheGluAsnGluPheLysLysAlaSer 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        507 GAAGTGGGTTTGTTTAATGAGTTGGCGAGTCCATTTGAGAATGAATTCAAGAAAGCTTCA 566

Qy        101 GluAspAspIleLysLysMetProLeuAspLeuSerProLeuAlaThrProIleIleArg 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        567 GAAGATGACATTAAAAAAATGCCTCTAGATTTATCCCCTCTTGCAACACCTATCATAAGA 626

Qy        121 SerLysIleGluGluProSerValValGluThrThrHisGlnAspSerProLeuProHis 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        627 AGCAAAATTGAGGAGCCTTCTGTTGTAGAAACAACTCACCAGGATAGTCCTTTACCTCAC 686

Qy        141 ProGluSerThrThrSerAspGluLysGluValProLeuAlaGlnThrAlaGlnProThr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        687 CCAGAGTCTACTACCAGTGATGAGAAGGAAGTACCATTGGCACAAACTGCACAGCCCACA 746

Qy        161 SerAlaIleValArgProAlaSerLeuGlnValProAsnValLeuLeuThrSerSerAsp 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        747 TCAGCTATTGTTCGTCCAGCATCATTACAGGTTCCCAATGTGCTGCTTACAAGTTCTGAC 806

Qy        181 SerSerValIleIleGlnGlnAlaValProSerProThrSerSerThrValIleThrGln 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        807 TCAAGTGTAATTATTCAGCAGGCAGTACCTTCACCAACCTCAAGTACTGTAATCACCCAG 866

Qy        201 AlaProSerSerAsnArgProIleVal 209
              ||||||||||||||||||||||||||||

Db        867 GCACCATCCTCTAACAGGCCAATTGTG 893

RESULT 2
AX821940
LOCUS       AX821940               1647 bp    DNA     linear   PAT 10-DEC-2003
DEFINITION  Sequence 68 from Patent WO03068961.
ACCESSION   AX821940
VERSION     AX821940.1  GI:39725161
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Andrews,P.A., Walsh,J.A. and Gokhale,P.A.
  TITLE     Method to modify differentiation of pluripotential stem cells
  JOURNAL   Patent: WO 03068961-A 68 21-AUG-2003;
            Axordia Limited (GB)
FEATURES             Location/Qualifiers
     source          1. .1647
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              2.31e-209       Length:         1647
Score:                  209.00          Matches:        209
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 8, 2005, 13:34:22 ; Search time 559 Seconds
(without alignments)
144.442 Million cell updates/sec

Title: AAH26175
Perfect score: 209
Sequence: 1 MKFKLHVNSARQYKDLWNMS.....PSPTSSTVITQAPSSNRPIV 209

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1726216 seqs, 386330316 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1726216

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep:*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	29	13.9	29	9	US-09-970-515-20	Sequence 20, Appl
2	29	13.9	29	14	US-10-165-250A-20	Sequence 20, Appl
3	29	13.9	29	15	US-10-342-683-20	Sequence 20, Appl
4	29	13.9	29	15	US-10-457-614A-20	Sequence 20, Appl
5	29	13.9	29	17	US-10-924-028-20	Sequence 20, Appl
6	27	12.9	27	13	US-10-115-178-6	Sequence 6, Appli
7	17	8.1	17	14	US-10-211-088-91	Sequence 91, Appl
8	16	7.7	17	14	US-10-211-088-79	Sequence 79, Appl
9	8	3.8	38	15	US-10-424-599-195921	Sequence 195921,
10	8	3.8	125	16	US-10-425-115-232009	Sequence 232009,
11	8	3.8	153	17	US-10-732-923-20872	Sequence 20872, A
12	8	3.8	523	16	US-10-437-963-129920	Sequence 129920,
13	8	3.8	951	16	US-10-437-963-167485	Sequence 167485,
14	8	3.8	1201	16	US-10-437-963-129457	Sequence 129457,
15	7	3.3	36	15	US-10-424-599-160301	Sequence 160301,
16	7	3.3	62	16	US-10-425-115-276081	Sequence 276081,
17	7	3.3	72	9	US-09-867-550-292	Sequence 292, App
18	7	3.3	74	15	US-10-424-599-256893	Sequence 256893,
19	7	3.3	78	15	US-10-424-599-212020	Sequence 212020,
20	7	3.3	79	15	US-10-424-599-145651	Sequence 145651,
21	7	3.3	100	16	US-10-425-115-276801	Sequence 276801,
22	7	3.3	101	16	US-10-425-115-285382	Sequence 285382,
23	7	3.3	101	16	US-10-425-115-351511	Sequence 351511,
24	7	3.3	115	16	US-10-425-115-298817	Sequence 298817,
25	7	3.3	120	9	US-09-738-626-6896	Sequence 6896, Ap
26	7	3.3	124	14	US-10-080-254-97	Sequence 97, Appl
27	7	3.3	124	15	US-10-242-355-641	Sequence 641, App
28	7	3.3	132	15	US-10-424-599-238800	Sequence 238800,
29	7	3.3	136	16	US-10-425-115-348784	Sequence 348784,
30	7	3.3	139	16	US-10-425-115-309456	Sequence 309456,
31	7	3.3	144	9	US-09-789-561-102	Sequence 102, App
32	7	3.3	144	11	US-09-833-245-2138	Sequence 2138, Ap
33	7	3.3	144	15	US-10-424-599-271504	Sequence 271504,
34	7	3.3	144	17	US-10-883-936-102	Sequence 102, App
35	7	3.3	147	16	US-10-425-115-320707	Sequence 320707,
36	7	3.3	147	16	US-10-425-115-358340	Sequence 358340,
37	7	3.3	151	15	US-10-424-599-257328	Sequence 257328,
38	7	3.3	153	15	US-10-425-114-66599	Sequence 66599, A
39	7	3.3	155	15	US-10-425-114-66627	Sequence 66627, A
40	7	3.3	159	15	US-10-425-114-63911	Sequence 63911, A
41	7	3.3	162	16	US-10-767-701-38653	Sequence 38653, A
42	7	3.3	164	16	US-10-437-963-198026	Sequence 198026,
43	7	3.3	174	9	US-09-839-185-12	Sequence 12, Appl
44	7	3.3	179	16	US-10-437-963-140436	Sequence 140436,
45	7	3.3	184	15	US-10-282-122A-52382	Sequence 52382, A

ALIGNMENTS

RESULT 1
US-09-970-515-20
; Sequence 20, Application US/09970515
; Patent No. US20020127676A1
; GENERAL INFORMATION:
; APPLICANT: Bonny, Christophe
; TITLE OF INVENTION: CELL-PERMEABLE PEPTIDE INHIBITORS OF THE JNK SIGNAL
; TITLE OF INVENTION: TRANSDUCTION PATHWAY
; FILE REFERENCE: 20349-501 DIV
; CURRENT APPLICATION NUMBER: US/09/970,515
; CURRENT FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: 09/503,954
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: USSN 60/158,774
; PRIOR FILING DATE: 1999-10-12
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically synthesized
US-09-970-515-20

Query Match 13.9%; Score 29; DB 9; Length 29;
Best Local Similarity 100.0%; Pred. No. 2.5e-20;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 37 TNEDHLAVHKHKHEMTLKFGPARNDSVIV 65
||||||||||||||||||||||||||||||
Db 1 TNEDHLAVHKHKHEMTLKFGPARNDSVIV 29

```
RESULT 2
US-10-165-250A-20
; Sequence 20, Application US/10165250A
; Publication No. US20030108539A1
; GENERAL INFORMATION:
;  APPLICANT: Bonny, Christophe
;  TITLE OF INVENTION: CELL-PERMEABLE PEPTIDE INHIBITORS OF THE JNK SIGNAL
;  TITLE OF INVENTION:  TRANSDUCTION PATHWAY
;  FILE REFERENCE: 20349-501 CIP
;  CURRENT APPLICATION NUMBER: US/10/165,250A
;  CURRENT FILING DATE:  2002-06-07
;  PRIOR APPLICATION NUMBER: 09/503,954
;  PRIOR FILING DATE: 2000-02-14
;  PRIOR APPLICATION NUMBER: USSN 60/158,774
;  PRIOR FILING DATE: 1999-10-12
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
;   LENGTH: 29
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically synthesized
US-10-165-250A-20

  Query Match             13.9%;  Score 29;  DB 14;  Length 29;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-20;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         37 TNEDHLAVHKHKHEMTLKFGPARNDSVIV 65
              |||||||||||||||||||||||||||||
Db          1 TNEDHLAVHKHKHEMTLKFGPARNDSVIV 29


RESULT 3
US-10-342-683-20
; Sequence 20, Application US/10342683
; Publication No. US20030220480A1
; GENERAL INFORMATION:
;  APPLICANT: Bonny, Christophe
;  TITLE OF INVENTION: CELL-PERMEABLE PEPTIDE INHIBITORS OF THE JNK SIGNAL
;  TITLE OF INVENTION:  TRANSDUCTION PATHWAY
;  FILE REFERENCE: 20349-501
;  CURRENT APPLICATION NUMBER: US/10/342,683
;  CURRENT FILING DATE:  2003-01-15
;  PRIOR APPLICATION NUMBER: US/09/503,954A
;  PRIOR FILING DATE: 2000-02-14
;  PRIOR APPLICATION NUMBER: USSN 60/158,774
;  PRIOR FILING DATE: 1999-10-12
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
;   LENGTH: 29
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically synthesized
US-10-342-683-20

  Query Match             13.9%;  Score 29;  DB 15;  Length 29;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-20;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         37 TNEDHLAVHKHKHEMTLKFGPARNDSVIV 65
              |||||||||||||||||||||||||||||
Db          1 TNEDHLAVHKHKHEMTLKFGPARNDSVIV 29


RESULT 4
US-10-457-614A-20
; Sequence 20, Application US/10457614A
; Publication No. US20040082509A1
; GENERAL INFORMATION:
;  APPLICANT: Bonny, Christope
;  TITLE OF INVENTION: CELL-PERMEABLE PEPTIDE INHIBITORS OF THE JNK SIGNAL
;  TITLE OF INVENTION:  TRANSDUCTION PATHWAY
;  FILE REFERENCE: 20349-501CIP2
;  CURRENT APPLICATION NUMBER: US/10/457,614A
;  CURRENT FILING DATE:  2003-06-09
;  PRIOR APPLICATION NUMBER: 60/158,774
;  PRIOR FILING DATE: 1999-10-12
;  PRIOR APPLICATION NUMBER: 09/503,954
;  PRIOR FILING DATE: 2000-02-14
;  PRIOR APPLICATION NUMBER: 10/165,250
;  PRIOR FILING DATE: 2002-06-07
;  PRIOR APPLICATION NUMBER: 60/347,062
;  PRIOR FILING DATE: 2002-01-09
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
;   LENGTH: 29
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically synthesized
US-10-457-614A-20

  Query Match             13.9%;  Score 29;  DB 15;  Length 29;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-20;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         37 TNEDHLAVHKHKHEMTLKFGPARNDSVIV 65
              |||||||||||||||||||||||||||||
Db          1 TNEDHLAVHKHKHEMTLKFGPARNDSVIV 29


RESULT 5
US-10-924-028-20
; Sequence 20, Application US/10924028
; Publication No. US20050043241A1
; GENERAL INFORMATION:
;  APPLICANT: Bonny, Christophe
;  TITLE OF INVENTION: CELL-PERMEABLE PEPTIDE INHIBITORS OF THE JNK SIGNAL
;  TITLE OF INVENTION:  TRANSDUCTION PATHWAY
;  FILE REFERENCE: 20349-501 DIV
;  CURRENT APPLICATION NUMBER: US/10/924,028
;  CURRENT FILING DATE:  2004-08-23
;  PRIOR APPLICATION NUMBER: US/09/970,515
;  PRIOR FILING DATE: 2001-10-03
;  PRIOR APPLICATION NUMBER: 09/503,954
;  PRIOR FILING DATE: 2000-02-14
;  PRIOR APPLICATION NUMBER: USSN 60/158,774
;  PRIOR FILING DATE: 1999-10-12
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
;   LENGTH: 29
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically synthesized
US-10-924-028-20

  Query Match             13.9%;  Score 29;  DB 17;  Length 29;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-20;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         37 TNEDHLAVHKHKHEMTLKFGPARNDSVIV 65
              |||||||||||||||||||||||||||||
Db          1 TNEDHLAVHKHKHEMTLKFGPARNDSVIV 29


RESULT 6
US-10-115-178-6
; Sequence 6, Application US/10115178
; Publication No. US20020119135A1
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 8, 2005, 13:22:35 ; Search time 30 Seconds
(without alignments)
520.055 Million cell updates/sec

Title: AAH26175
Perfect score: 209
Sequence: 1 MKFKLHVNSARQYKDLWNMS.....PSPTSSTVITQAPSSNRPIV 209

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 513545 seqs, 74649064 residues

Word size : 0

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	209	100.0	501	4	US-09-949-016-11281	Sequence 11281, A
2	31	14.8	52	4	US-09-513-999C-7023	Sequence 7023, Ap
3	29	13.9	29	4	US-09-503-954A-20	Sequence 20, Appl
4	29	13.9	29	4	US-09-970-515A-20	Sequence 20, Appl
5	27	12.9	27	3	US-08-819-177-6	Sequence 6, Appli
6	8	3.8	1053	4	US-09-328-352-5058	Sequence 5058, Ap
7	7	3.3	22	3	US-08-256-747C-1	Sequence 1, Appli
8	7	3.3	22	3	US-08-834-130A-1	Sequence 1, Appli
9	7	3.3	25	3	US-08-256-747C-57	Sequence 57, Appl
10	7	3.3	25	3	US-08-834-130A-57	Sequence 57, Appl
11	7	3.3	96	4	US-09-513-999C-7523	Sequence 7523, Ap
12	7	3.3	134	3	US-09-605-858-33	Sequence 33, Appl
13	7	3.3	158	4	US-09-902-540-11525	Sequence 11525, A
14	7	3.3	193	4	US-09-248-796A-16279	Sequence 16279, A
15	7	3.3	196	4	US-09-949-016-11091	Sequence 11091, A
16	7	3.3	225	4	US-09-902-540-9773	Sequence 9773, Ap
17	7	3.3	229	5	PCT-US96-03916-13	Sequence 13, Appl
18	7	3.3	229	5	PCT-US96-03916-62	Sequence 62, Appl
19	7	3.3	258	4	US-09-485-529-20	Sequence 20, Appl
20	7	3.3	396	3	US-09-134-001C-4546	Sequence 4546, Ap
21	7	3.3	405	4	US-09-248-796A-17965	Sequence 17965, A
22	7	3.3	418	4	US-09-710-279-380	Sequence 380, App
23	7	3.3	430	3	US-09-134-001C-4302	Sequence 4302, Ap
24	7	3.3	447	4	US-09-248-796A-20950	Sequence 20950, A
25	7	3.3	481	1	US-07-927-071-1	Sequence 1, Appli
26	7	3.3	481	6	5164481-1	Patent No. 5164481
27	7	3.3	481	6	5164481-1	Patent No. 5164481
28	7	3.3	690	4	US-09-538-092-1176	Sequence 1176, Ap
29	7	3.3	714	4	US-09-248-796A-19419	Sequence 19419, A
30	7	3.3	859	4	US-09-538-092-717	Sequence 717, App
31	7	3.3	992	1	US-08-127-499A-1	Sequence 1, Appli
32	7	3.3	992	1	US-08-482-847-1	Sequence 1, Appli
33	7	3.3	1063	1	US-08-093-453B-3	Sequence 3, Appli
34	7	3.3	1063	1	US-08-127-499A-8	Sequence 8, Appli
35	7	3.3	1063	1	US-08-482-847-8	Sequence 8, Appli
36	6	2.9	10	5	PCT-US95-03236-11	Sequence 11, Appl
37	6	2.9	17	3	US-09-308-022-3	Sequence 3, Appli
38	6	2.9	18	2	US-08-480-190-204	Sequence 204, App
39	6	2.9	18	2	US-08-488-379-204	Sequence 204, App
40	6	2.9	18	4	US-08-475-399A-204	Sequence 204, App
41	6	2.9	18	4	US-08-077-255A-204	Sequence 204, App
42	6	2.9	18	5	PCT-US93-07545-204	Sequence 204, App
43	6	2.9	20	2	US-08-934-915-8	Sequence 8, Appli
44	6	2.9	20	3	US-09-000-094-48	Sequence 48, Appl
45	6	2.9	20	4	US-10-011-749-48	Sequence 48, Appl

ALIGNMENTS

RESULT 1
US-09-949-016-11281
; Sequence 11281, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11281
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11281

Query Match 100.0%; Score 209; DB 4; Length 501;
Best Local Similarity 100.0%; Pred. No. 1.6e-193;
Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MKFKLHVNSARQYKDLWNMSDDKPFLCTAPGCGQRFTNEDHLAVHKHKHEMTLKFGPARN 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          5 MKFKLHVNSARQYKDLWNMSDDKPFLCTAPGCGQRFTNEDHLAVHKHKHEMTLKFGPARN 64

Qy         61 DSVIVADQTPTPTRFLKNCEEVGLFNELASPFENEFKKASEDDIKKMPLDLSPLATPIIR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 DSVIVADQTPTPTRFLKNCEEVGLFNELASPFENEFKKASEDDIKKMPLDLSPLATPIIR 124

Qy        121 SKIEEPSVVETTHQDSPLPHPESTTSDEKEVPLAQTAQPTSAIVRPASLQVPNVLLTSSD 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 SKIEEPSVVETTHQDSPLPHPESTTSDEKEVPLAQTAQPTSAIVRPASLQVPNVLLTSSD 184

Qy        181 SSVIIQQAVPSPTSSTVITQAPSSNRPIV 209
              ||||||||||||||||||||||||||||||
Db        185 SSVIIQQAVPSPTSSTVITQAPSSNRPIV 213
```

RESULT 2
US-09-513-999C-7023
; Sequence 7023, Application US/09513999C
; Patent No. 6783961

```
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 7023
;   LENGTH: 52
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: 6
;   OTHER INFORMATION: Xaa=Pro or Ser
US-09-513-999C-7023

  Query Match             14.8%;  Score 31;  DB 4;  Length 52;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-22;
  Matches   31;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         29 APGCGQRFTNEDHLAVHKHKHEMTLKFGPAR 59
              |||||||||||||||||||||||||||||||
Db         11 APGCGQRFTNEDHLAVHKHKHEMTLKFGPAR 41


RESULT 3
US-09-503-954A-20
; Sequence 20, Application US/09503954A
; Patent No. 6610820
; GENERAL INFORMATION:
;  APPLICANT: Bonny, Christophe
;  TITLE OF INVENTION: CELL-PERMEABLE PEPTIDE INHIBITORS OF THE JNK SIGNAL
;  TITLE OF INVENTION:  TRANSDUCTION PATHWAY
;  FILE REFERENCE: 20349-501
;  CURRENT APPLICATION NUMBER: US/09/503,954A
;  CURRENT FILING DATE:  2000-02-14
;  PRIOR APPLICATION NUMBER: USSN 60/158,774
;  PRIOR FILING DATE: 1999-10-12
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
;   LENGTH: 29
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically synthesized
US-09-503-954A-20

  Query Match             13.9%;  Score 29;  DB 4;  Length 29;
  Best Local Similarity   100.0%;  Pred. No. 7e-21;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         37 TNEDHLAVHKHKHEMTLKFGPARNDSVIV 65
              |||||||||||||||||||||||||||||
Db          1 TNEDHLAVHKHKHEMTLKFGPARNDSVIV 29


RESULT 4
US-09-970-515A-20
; Sequence 20, Application US/09970515A
; Patent No. 6780970
; GENERAL INFORMATION:
;  APPLICANT: Bonny, Christophe
;  TITLE OF INVENTION: CELL-PERMEABLE PEPTIDE INHIBITORS OF THE JNK SIGNAL
;  TITLE OF INVENTION:  TRANSDUCTION PATHWAY
;  FILE REFERENCE: 20349-501 DIV
;  CURRENT APPLICATION NUMBER: US/09/970,515A
;  CURRENT FILING DATE:  2001-10-03
;  PRIOR APPLICATION NUMBER: 09/503,954
;  PRIOR FILING DATE: 2000-02-14
;  PRIOR APPLICATION NUMBER: 60/158,774
;  PRIOR FILING DATE: 1999-10-12
;  NUMBER OF SEQ ID NOS: 21
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
;   LENGTH: 29
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically synthesized
US-09-970-515A-20

  Query Match             13.9%;  Score 29;  DB 4;  Length 29;
  Best Local Similarity   100.0%;  Pred. No. 7e-21;
  Matches   29;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         37 TNEDHLAVHKHKHEMTLKFGPARNDSVIV 65
              |||||||||||||||||||||||||||||
Db          1 TNEDHLAVHKHKHEMTLKFGPARNDSVIV 29


RESULT 5
US-08-819-177-6
; Sequence 6, Application US/08819177
; Patent No. 6043083
;  GENERAL INFORMATION:
;    APPLICANT:  Davis, Roger J.
;    APPLICANT:  Dickens, Martin
;    TITLE OF INVENTION:  INHIBITORS OF THE JNK SIGNAL
;    TITLE OF INVENTION:  TRANSDUCTION PATHWAY AND METHODS OF USE
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson, P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows95
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/819,177
;      FILING DATE:  28 April 1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fasse, J. Peter
;      REGISTRATION NUMBER:  32,983
;      REFERENCE/DOCKET NUMBER:  07917/037001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-542-5070
;      TELEFAX:  617-542-8906
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  27 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      OTHER INFORMATION:  ATF-2 JNK-binding domain
US-08-819-177-6

  Query Match             12.9%;  Score 27;  DB 3;  Length 27;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-19;
  Matches   27;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 8, 2005, 11:16:22 ; Search time 92 Seconds
(without alignments)
878.619 Million cell updates/sec

Title: AAH26175
Perfect score: 209
Sequence: 1 MKFKLHVNSARQYKDLWNMS.....PSPTSSTVITQAPSSNRPIV 209

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 2105692 seqs, 386760381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	144	68.9	482	8	ADR10403	Adr10403 Human pro
2	58	27.8	411	5	ABB57020	Abb57020 Mouse isc
3	31	14.8	52	3	AAG02942	Aag02942 Human sec
4	31	14.8	413	8	ADP80860	Adp80860 Mouse C13
5	29	13.9	29	4	AAB68325	Aab68325 c-Jun ami
6	29	13.9	29	7	ADF17925	Adf17925 Synthetic
7	29	13.9	29	8	ADO01349	Ado01349 c-Jun ami
8	19	9.1	501	7	ADC18733	Adc18733 Human cyc
9	17	8.1	17	7	ADC22242	Adc22242 Protein b
10	16	7.7	17	7	ADC22230	Adc22230 Protein b
11	9	4.3	359	4	ABB67482	Abb67482 Drosophil
12	8	3.8	95	4	AAG78911	Aag78911 Human pro
13	8	3.8	835	3	AAG31412	Aag31412 Arabidops
14	8	3.8	929	3	AAG31411	Aag31411 Arabidops
15	8	3.8	937	3	AAG31410	Aag31410 Arabidops
16	8	3.8	937	6	ABP59344	Abp59344 AT4G10590
17	8	3.8	1053	6	ADA33771	Ada33771 Acinetoba
18	7	3.3	22	2	AAR39825	Aar39825 E1 peptid
19	7	3.3	25	2	AAR39879	Aar39879 Lipopepti
20	7	3.3	49	4	AAU31646	Aau31646 Novel hum
21	7	3.3	67	4	AAM91309	Aam91309 Human imm
22	7	3.3	72	5	ABP63776	Abp63776 Human ORF
23	7	3.3	83	4	AAO13691	Aao13691 Human pol
24	7	3.3	91	4	ABB10659	Abb10659 Human pan
25	7	3.3	91	4	AAM92701	Aam92701 Human dig
26	7	3.3	96	3	AAG03442	Aag03442 Human sec
27	7	3.3	99	4	AAU48473	Aau48473 Propionib
28	7	3.3	99	6	ABM44992	Abm44992 Propionib
29	7	3.3	103	2	AAR29888	Aar29888 HCV NS4-N
30	7	3.3	114	8	ADS10644	Ads10644 Human the
31	7	3.3	120	4	AAG93142	Aag93142 C glutami
32	7	3.3	124	4	AAM99904	Aam99904 Human exc
33	7	3.3	124	4	AAM43703	Aam43703 Human bla
34	7	3.3	124	8	ADF71662	Adf71662 Human bla
35	7	3.3	134	4	AAB61536	Aab61536 DegP prot
36	7	3.3	135	4	AAU45096	Aau45096 Propionib
37	7	3.3	135	6	ABM41615	Abm41615 Propionib
38	7	3.3	144	4	AAB87361	Aab87361 Human gen
39	7	3.3	144	5	ABG65389	Abg65389 Human alb
40	7	3.3	144	6	ABP58866	Abp58866 Tetrapept
41	7	3.3	144	8	ADL78656	Adl78656 Albumin f
42	7	3.3	158	8	ADN47355	Adn47355 Thermococ
43	7	3.3	165	4	AAO04300	Aao04300 Human pol
44	7	3.3	168	5	ABB97806	Abb97806 Human sec
45	7	3.3	168	6	ABP75874	Abp75874 Human sec

ALIGNMENTS

RESULT 1
ADR10403
ID ADR10403 standard; protein; 482 AA.
XX
AC ADR10403;
XX
DT 04-NOV-2004 (first entry)
XX
DE Human protein useful for treating neurological disease Seq 3909.
XX
KW human; oligo-capping method; diagnostic marker; gene therapy;
KW osteoporosis; neurological disease; Alzheimer's disease;
KW Parkinson's disease; dementia; short memory; cancer;
KW sense or motor function; emotional reaction; fear response; panic;
KW osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;
KW tranquiliser.
XX
OS Homo sapiens.
XX
PN EP1447413-A2.
XX
PD 18-AUG-2004.
XX
PF 12-FEB-2004; 2004EP-00003145.
XX
PR 14-FEB-2003; 2003JP-00102207.
PR 09-MAY-2003; 2003JP-00131452.
XX
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Isogai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T;
PI Wakamatsu A, Ishii S, Nagai K, Irie R;
XX
DR WPI; 2004-583265/57.
DR N-PSDB; ADR08447.
XX
PT New 1995 cDNA, useful for treating osteoporosis, neurological diseases,
PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX
PS Claim 1; SEQ ID NO 3909; 2686pp; English.
XX
CC This invention relates to novel, isolated full length human cDNA
CC molecules and the encoded proteins thereof. Specifically, it refers to
CC cDNA clones obtained by an oligo-capping method, where none of these
CC clones are identical to any known human mRNAs. The present invention
CC describes an immunoassay to identify agonists and antagonists, as well as
CC antibodies, antisense molecules and siRNAs that can all be used to bind
CC to and modulate expression of the cDNA molecules. As such, these

CC molecules are useful for diagnostic markers or therapeutic targets for
CC the various diseases or morbid states. In particular, they are useful in
CC gene therapy for treating osteoporosis, neurological disease, Alzheimer's
CC disease, Parkinson's disease, dementia, short memory and various cancers,
CC as well as for maintaining equilibrium of sense or motor function, and
CC for treating emotional reaction, fear response and panic. Accordingly,
CC they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian,
CC cytostatic and tranquiliser activities. This polypeptide is a protein
CC encoded by a full length human cDNA sequence of the invention. NOTE: This
CC sequence is not given in the sequence listing of the specification but
CC can be obtained on CD-ROM from the European Patent Office, Vienna Sub-
CC office.
XX
SQ Sequence 482 AA;

Query Match 68.9%; Score 144; DB 8; Length 482;
Best Local Similarity 100.0%; Pred. No. 2e-137;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 66 ADQTPTPTRFLKNCEEVGLFNELASPFENEFKKASEDDIKKMPLDLSPLATPIIRSKIEE 125
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 43 ADQTPTPTRFLKNCEEVGLFNELASPFENEFKKASEDDIKKMPLDLSPLATPIIRSKIEE 102

Qy 126 PSVVETTHQDSPLPHPESTTSDEKEVPLAQTAQPTSAIVRPASLQVPNVLLTSSDSSVII 185
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 103 PSVVETTHQDSPLPHPESTTSDEKEVPLAQTAQPTSAIVRPASLQVPNVLLTSSDSSVII 162

Qy 186 QQAVPSPTSSTVITQAPSSNRPIV 209
|||||||||||||||||||||||||
Db 163 QQAVPSPTSSTVITQAPSSNRPIV 186

RESULT 2
ABB57020
ID ABB57020 standard; protein; 411 AA.
XX
AC ABB57020;
XX
DT 07-MAR-2002 (first entry)
XX
DE Mouse ischaemic condition related protein sequence SEQ ID NO:2.
XX
KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
KW vasospastic ischaemia; ischaemic condition; ischaemic disease.
XX
OS Mus musculus.
XX
PN WO200188188-A2.
XX
PD 22-NOV-2001.
XX
PF 18-MAY-2001; 2001WO-JP004192.
XX
PR 18-MAY-2000; 2000JP-00145977.
XX
PA (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX
PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
XX
DR WPI; 2002-034733/04.
DR N-PSDB; ABI99202.
XX
PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT expression levels of particular genes defined in the specification or by
PT determining the expression profile of a gene group comprising these
PT genes.
XX
PS Claim 2; Page 51-53; 2690pp; English.
XX
CC The present invention describes a method for examining ischaemic
CC conditions, comprising measuring the expression levels of particular
CC genes (I) in a test sample or determining the expression profile of a
CC gene group in the sample comprising genes selected from (I). The method
CC is useful for examining the ischaemic condition (e.g. compressive
CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
CC expression levels of particular genes (ABI99202 to ABI99912, encoding the
CC protein sequences in ABB57020 to ABB57374) or by determining the
CC expression profile of a gene group comprising these genes. The expression
CC levels or expression profiles produced by these genes are used as an
CC indicator when screening for ischaemic condition-improving drugs or
CC therapeutics for ischaemic diseases. ABI99913 and ABI99914 represent PCR
CC primers for a mouse ischaemic condition related sequence, which are used
CC in the exemplification of the present invention
XX
SQ Sequence 411 AA;

Query Match 27.8%; Score 58; DB 5; Length 411;
Best Local Similarity 100.0%; Pred. No. 5.5e-50;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 152 PLAQTAQPTSAIVRPASLQVPNVLLTSSDSSVIIQQAVPSPTSSTVITQAPSSNRPIV 209
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 58 PLAQTAQPTSAIVRPASLQVPNVLLTSSDSSVIIQQAVPSPTSSTVITQAPSSNRPIV 115

RESULT 3
AAG02942
ID AAG02942 standard; protein; 52 AA.
XX
AC AAG02942;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein, SEQ ID NO: 7023.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-00200610.
XX
PR 26-FEB-1999; 99US-0122487P.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI; 2000-500381/45.
DR N-PSDB; AAC02948.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
PS Claim 13; SEQ ID NO 7023; 71pp + Sequence Listing; English.
XX
CC The present sequence is a polypeptide encoded by one of a large number of
CC 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs were
CC prepared from total human RNAs or polyA+ RNAs derived from 30 different
CC tissues. EST sequences usually correspond mainly to the 3' untranslated
CC region (UTR) of the mRNA because they are often obtained from oligo-dT
CC primed cDNA libraries. Such ESTs are not well suited for isolating cDNA
CC sequences derived from the 5' ends of mRNAs and even in those cases where
CC longer cDNA sequences have been obtained, the full 5' UTR is rarely
CC included. 5' ESTs are derived from mRNAs with intact 5' ends and can
CC therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs
CC are also used in diagnostic, forensic, gene therapy and chromosome
CC mapping procedures. They are used to obtain upstream regulatory sequences
CC and to design expression and secretion vectors
XX
SQ Sequence 52 AA;